



# Simulation du mouvement brownien et des diffusions

Olivier Faure

## ► To cite this version:

Olivier Faure. Simulation du mouvement brownien et des diffusions. Sciences de la Terre. Ecole Nationale des Ponts et Chaussées, 1992. Français. NNT: . tel-00523258

**HAL Id: tel-00523258**

**<https://pastel.archives-ouvertes.fr/tel-00523258>**

Submitted on 4 Oct 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

X

Ecole Nationale des Ponts et Chaussées

## Thèse de doctorat

Spécialité

Mathématiques appliquées.



Sujet de thèse:

*Simulation du mouvement brownien et des diffusions.*

Présentée par Olivier Faure .

et soutenue le 21 Février 1992

devant le jury composé de:

Nicolas Bouleau (Président)

Denis Talay (Rapporteur)

Nigel Newton (Rapporteur)

J.M.C. Clark

Bernard Lapeyre

Luciano Tubaro



DOC08389



## Avant propos

Commençons par quelques réflexions, mûries tout au long de ce travail, et dont l'objet est surtout de situer cette thèse dans son contexte.

Entreprendre un doctorat, quand on est ingénieur, passe parfois encore pour une fantaisie. On peut pourtant s'étonner que la formation par la recherche, si prisée chez nos voisins allemands ou américains, ne le soit pas davantage chez nous. Gageons que cette attitude changera. Il m'apparaît en effet *a posteriori* que cette aventure est valorisante à beaucoup de points de vue; c'est notamment l'occasion d'acquérir un corpus de connaissances solides dans un domaine un peu "pointu", qui complète la formation plutôt généraliste de nos écoles. C'est aussi l'occasion d'avoir une attitude scientifique *active*, en créant quelque chose, aussi modeste soit cette création.

Si j'ai choisi le domaine des probabilités, je le dois avant tout à la personnalité de mon directeur de thèse. Il a su, par son enseignement et ses ouvrages, me donner goût aux probabilités puis éveiller ma curiosité au point de m'inciter à aller plus loin pour la satisfaire. Et puis quoi de plus naturel que cette fascination pour le hasard que les joueurs de pokers ou de roulette connaissent bien, et donc ce désir de comprendre et maîtriser ce hasard. Et pourtant.

J'ai souvent constaté chez mes interlocuteurs de l'incrédulité, voire de la suspicion, quand j'affirmais préparer une thèse en probabilités. Il n'était pas facile de convaincre que les probabilistes ne passent pas leur temps à tirer des boules de toutes les couleurs d'une urne anonyme, avec ou sans remise, à des fins statistiques. Et, tentant d'expliquer et d'illustrer ce qu'est un processus stochastique par l'exemple simple du mouvement brownien, je ne laissais bien souvent que le souvenir d'un botaniste excentrique observant une colonie anarchique de grains de pollen en suspension dans l'eau. La crédibilité de mon propos était parfois rétablie en citant l'utilisation récente du calcul stochastique en finance pour fixer le juste prix d'une option; preuve irréfutable de "l'utilité" d'un tel calcul...

Au delà de cette petite histoire, il faut bien reconnaître qu'il y a un décalage surprenant entre l'énorme littérature théorique consacrée aux probabilités, réservée à quelques initiés, et l'utilisation qui est faite de cette littérature dans les applications. Ce décalage, sans doute en partie responsable de cette perception réductrice des probabilités, est d'autant plus surprenant que ces dernières sont nées par la modélisation des jeux de hasard, et donc par la pratique. Essayons d'indiquer, puisque tel est au fond l'esprit de cette thèse, quels rôles la *simulation numérique* peut jouer dans cette histoire.

Bien sûr elle est le maillon indispensable entre théorie et applications: la plupart du temps, ne connaissant pas de solution analytique simple au problème étudié, simuler permet une résolution quasi-expérimentale. La simulation peut aussi avoir un rôle plus interne aux probabilités, celle

d'un outils d'expérimentation permettant de confirmer, d'infirmer, ou de découvrir de nouvelles propriétés mathématiques. Mais son rôle peut aller plus loin. Une des difficultés d'apprentissage (ou de vulgarisation) des probabilités me semble être *l'abondance de formalisme*. J'ai constamment buté sur ces  $(\Omega, \mathcal{A}, \mathbf{P}, \mathcal{F}_t)$ ,  $\mathbf{P}^x(W_t \in da, L(t) \in db, \Gamma_+(t) \in d\tau)$ ... en me demandant à quoi diable pouvait bien servir ce drôle de langage ésotérique, puis ce qu'il fallait comprendre par ces formules et enfin comment tout cela avait été inventé. Il est en réalité clair qu'un tel formalisme est naturel et nécessaire pour raisonner, et pour raisonner juste. Mais, s'il est vrai que la clef d'une bonne compréhension des probabilités repose sur une bonne intuition, alors la simulation numérique a aussi un rôle essentiel à jouer comme *stimulateur* d'intuition.

---

L'objet de cette thèse est l'étude de la simulation numérique de certains processus stochastiques, les diffusions, dont le mouvement brownien est un exemple typique.

Nous commençons par quelques rappels sur le mouvement brownien au chapitre 1. Il s'agit d'une présentation élémentaire, qui s'appuie sur la simulation numérique, et permet de rappeler quelques propriétés classiques. Puis nous présentons au chapitre 2 une simulation alternative du mouvement brownien, en un sens plus naturelle, qui s'attache davantage à son comportement spatial que les méthodes traditionnelles. Le mouvement brownien est simulé à des instants aléatoires qui gouvernent son comportement; ce sont les temps de sortie de certaines "boîtes noires". En choisissant la taille et la position de ces boîtes noires dans l'espace, et sous réserve qu'elles se chevauchent, on peut ainsi simuler très précisément une trajectoire brownienne.

La suite de la thèse est consacrée à l'analyse numérique des équations différentielles stochastiques (E.D.S) et à la simulation informatique de leur solution.

Nous commençons au chapitre 3 par une introduction qui rappelle ce que sont les E.D.S, cite quelques unes de leurs propriétés et applications classiques dans les sciences de l'ingénieur.

Au chapitre 4 nous présentons un résultat de convergence trajectorielle du schéma d'Euler en précisant l'ordre de convergence. Un résultat similaire est présenté pour le schéma de Milshtein au chapitre 5. Comme on peut s'y attendre, ce schéma est plus performant que le schéma d'Euler, quand la condition classique de commutativité est vérifiée. Ceci améliore partiellement un résultat de Denis Talay.

On étudie ensuite au chapitre 6 une classe de schémas de discrétisation à pas variables permettant une approximation *spatiale* des diffusions dans l'esprit du chapitre 2. Nous commençons par un résultat assez général de convergence d'un schéma d'Euler défini le long d'une subdivision aléatoire. Dans le cas où cette subdivision est gouvernée par les temps de passage successifs du mouvement brownien, nous retrouvons et étendons partiellement des travaux de Nigel Newton. Dans le cas où cette subdivision de façon à ce que les accroissements du schéma de discrétisation soient constants, nous étudions un schéma de discrétisation originalement présenté par Bichteler. Nous précisons sa vitesse de convergence et donnons une méthode de simulation numérique.

Le chapitre 7 est un panorama des travaux existants sur la discrétisation des équations

différentielles stochastiques. Sans prétendre être exhaustif, nous présentons au contraire une relecture des travaux existants dans l'optique de la simulation numérique.

Enfin le chapitre 8 s'attache à quelques questions ou problèmes non résolus qui représentent un intérêt évident pour les applications. Nous suggérons pour chacune de ces questions quelques commencements de réponse.

---

La démarche, ou l'esprit de ce travail, est donc de promouvoir la simulation numérique en probabilités. Comme nous avons essayé de l'expliquer, la motivation d'un tel travail est triple: vulgariser des objets probabilistes difficiles, faire un petit pas vers l'application des probabilités aux sciences dites de l'ingénieur, et enfin contribuer, modestement, à l'étude expérimentale des processus stochastiques.

Quiconque connaît Nicolas Bouleau, la diversité et l'originalité de ses travaux, aura reconnu sa profonde influence dans mes propos; cet esprit dont je parlais précédemment transparait dans ses ouvrages et publications ainsi que dans les travaux du CERMA. Il a eu la lourde tâche de diriger ma thèse. Plutôt que diriger, je devrais dire écouter, conseiller et orienter, tant il est vrai qu'il m'a laissé une grande liberté dans mon travail. Je tiens à lui exprimer ici ma profonde gratitude et à le remercier très chaleureusement.

Denis Talay a été sans le savoir un modèle. C'est tout d'abord sur sa thèse et ses publications que j'ai commencé à travailler. Mais c'est aussi ce côté jeune chercheur brillant et sympathique qui m'a poussé à continuer. Il a par ailleurs su m'encourager dans mon travail, et ce, de nombreuses fois. Je suis donc très heureux et reconnaissant qu'il ait bien voulu accepter la tâche ingrate de rapporteur. Je l'en remercie vivement.

C'est aussi sur les publications de Nigel Newton que j'ai planché et trouvé de l'inspiration. Je suis donc enchanté qu'il ait accepté d'être rapporteur et le remercie chaleureusement. C'est une tâche d'autant plus difficile pour lui que cette thèse est rédigée en français... et j'espère ne pas l'avoir dégoûté à jamais d'apprendre notre belle langue.

Bernard Lapeyre est plus particulièrement responsable de ce que l'on ne voit pas dans ces pages. C'est grâce à lui que j'ai appris à me servir d'une station Sun, d'Unix, de C, Latex, gnuplot... et en particulier pu donner corps à ce mémoire. Je le remercie de tant de généreux efforts et suis flatté qu'il fasse parti de ce jury.

C'est grâce à Martin Clark que j'ai pu entreprendre et mener à bout la rédaction de cette thèse; par le financement qu'il m'a obtenu au sein de l'Imperial College et l'accueil chaleureux qu'il m'a réservé. Je continue par ailleurs à travailler sous sa direction sur des sujets voisins. Je suis donc très heureux qu'il ait bien voulu prendre part à ce jury et le remercie vivement.

Luciano Tubaro, qui possède des talents d'analyste numéricien et de probabiliste, a accepté de participer à ce jury. Je suis très flatté qu'il soit venu de si loin et l'en remercie chaleureusement.

Damien Lamberton m'a toujours prêté une oreille attentive. Je l'ai de si nombreuses fois importuné par des questions triviales, il m'a de si nombreuses fois patiemment répondu, que je lui présente mes excuses et mes remerciements.

Je dois aussi remercier l'ensemble du CERMA, les autres chercheurs que j'ai aussi importunés par de triviales questions et qui ont bien voulu me répondre, les thésards avec qui j'ai partagé les mêmes doutes et les mêmes questions, et Véronique Serre qui contribue à faire de ce centre un lieu de travail convivial.

Enfin je remercie des êtres chers, qui se reconnaîtront dans ces lignes, et savent l'importance qu'a eu et a encore leur amitié ou leur amour.

*Un jour viendra que par une étude suivie de plusieurs siècles,  
les choses actuellement cachées paraîtront avec évidence,  
et la postérité s'étonnera que des vérités si claires nous aient échappé.  
( Sénèque )*

# Table des matières

<b>1</b>	<b>Simulation du mouvement brownien: Introduction</b>	<b>3</b>
<b>2</b>	<b>Simulation spatialement contrôlée du brownien</b>	<b>13</b>
2.1	Introduction.	14
2.2	Etude du temps de sortie $\tau_a$ d'un cube, dimension 1.	14
2.2.1	Quelques propriétés de $\tau_a$ .	14
2.2.2	Simulation de $(\tau_a, B_{\tau_a})$ .	16
2.3	Etude du temps de sortie $\tau_a$ d'un cube, dimension d.	21
2.3.1	Un algorithme de simulation de $(\tau_a, B_{\tau_a})$ .	21
2.4	Etude du temps de sortie $\tau_{a,\mu}$ d'un cube pour un brownien de drift $\mu$ .	25
2.4.1	Quelques propriétés de $\tau_{a,\mu}$ et $T_{a,\mu}$ .	26
2.4.2	Simulation de $(\tau_{a,\mu}, B_{\tau_{a,\mu}})$ .	29
2.4.3	Cas de la dimension d.	30
2.5	Simulation du brownien à sa sortie d'un ouvert borné.	31
2.5.1	Une méthode élémentaire de simulation.	32
2.5.2	La méthode des "boîtes de contrôle".	34
2.5.3	Application numérique au temps de sortie d'un cercle.	37
2.6	Conclusion.	41
<b>3</b>	<b>Simulation des diffusions: Introduction.</b>	<b>42</b>
3.1	Qu'est ce qu'une équation différentielle stochastique?	43
3.1.1	Un exemple introductif.	43
3.1.2	Propriétés classiques des diffusions.	44
3.1.3	Modélisation Itô ou Stratonovich ?	45
3.2	Qu'est ce qu'un schéma de discrétisation ?	48
3.2.1	Pourquoi discrétiser les E.D.S. ?	48
3.2.2	Définitions et notations.	49
3.3	Plan des chapitres sur la simulation des diffusions.	51
<b>4</b>	<b>Etude du schéma d'Euler.</b>	<b>52</b>
4.1	Introduction.	53
4.2	Convergence $L^p$ et $p.s$ du schéma	54



4.3	Résultats numériques . . . . .	61
4.4	Approximation du flot de l'E.D.S . . . . .	64
4.5	Convergence en loi du schéma. . . . .	65
<b>5</b>	<b>Etude du schéma de Milshtein.</b>	<b>68</b>
5.1	Introduction . . . . .	69
5.2	Cas de la dimension 1. . . . .	70
5.3	Cas de la dimension $n$ . . . . .	78
5.4	Liens avec l'E.D.S au sens Stratonovich. . . . .	80
<b>6</b>	<b>Des schémas de discrétisation à pas adaptés.</b>	<b>82</b>
6.1	Un résultat général de convergence. . . . .	83
6.1.1	Notations et hypothèses. . . . .	83
6.1.2	Résultat de convergence. . . . .	84
6.2	Un schéma avec contrôle spatial du brownien . . . . .	89
6.2.1	Résultat de convergence. . . . .	89
6.2.2	Simulation numérique: Un exemple d'application. . . . .	92
6.3	Un schéma avec contrôle spatial de la diffusion. . . . .	94
6.3.1	Résultat de convergence. . . . .	94
6.3.2	Simulation numérique: Un exemple d'application. . . . .	96
6.4	Quelques simulations numériques . . . . .	98
<b>7</b>	<b>Panorama des travaux existants.</b>	<b>101</b>
7.1	Introduction . . . . .	102
7.2	Approximations fortes. . . . .	103
7.2.1	Convergence quadratique. . . . .	103
7.2.2	Convergence p.s ou trajectorielle. . . . .	108
7.2.3	Le cas d'une E.D.S au sens Stratonovich. . . . .	109
7.3	Approximations en loi. . . . .	110
7.4	Autres types de convergence. . . . .	113
7.4.1	Calcul approché d'exposants de Lyapunov. . . . .	113
7.4.2	Approximation de la loi invariante d'un processus ergodique. . . . .	114
<b>8</b>	<b>Quelques problèmes ouverts.</b>	<b>115</b>
8.1	Approximation trajectorielle, cas non commutatif. . . . .	116
8.2	Stabilité des schémas de discrétisation. . . . .	117
8.3	Utilisation de pas variables pour l'approximation trajectorielle. . . . .	118
8.4	Sur l'approximation en loi . . . . .	119
<b>9</b>	<b>Bibliographie</b>	<b>120</b>
<b>10</b>	<b>Annexe I: Générateur pseudo-aléatoire et simulation du hasard.</b>	<b>127</b>

## **Chapitre 1**

# **Simulation du mouvement brownien: Introduction**

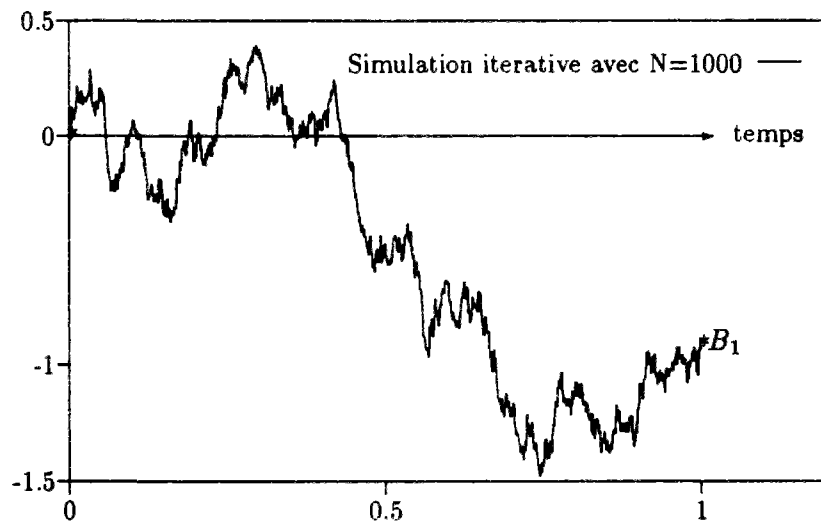
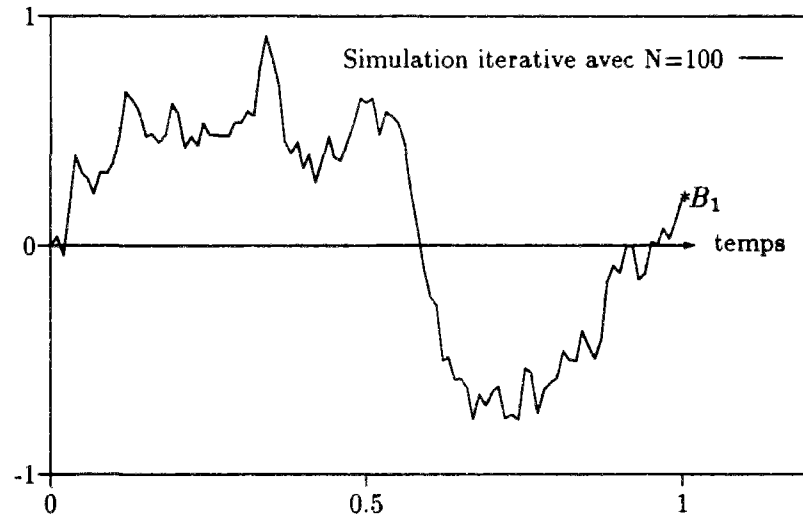
Le mouvement brownien intervient dans de très nombreuses applications des sciences de l'ingénieur, aussi bien que dans des phénomènes physiques très variés, comme nous le verrons au chapitre 3. Nous reviendrons plus en détail sur son utilisation comme modèle de "bruit blanc", mais disons qu'il est à la base de la modélisation de phénomènes bruités. Aussi allons-nous dans les deux chapitres suivants étudier le mouvement brownien d'un peu près, en mettant en place quelques outils de simulation numérique et quelques propriétés qui nous seront utiles par la suite.

C'est le botaniste anglais Brown qui le premier observe, en 1828, à l'aide d'un modeste microscope, des grains de pollen en suspension dans l'eau, et en décrit le mouvement extrêmement chaotique qui porte maintenant son nom. Puis Bachelier, Einstein et Langevin au tout début du siècle donnent les premières propriétés de ce mouvement; les accroissements ne sont pas corrélés et sont de loi gaussienne. Il faudra attendre N. Wiener en 1923, après l'apparition de la théorie de la mesure, pour en autre formaliser l'intuition de Bachelier selon laquelle ce mouvement chaotique est continu. P. Lévy donnera ensuite en 1939 et 1948 une autre construction du mouvement brownien et une description très profonde de sa structure fine. Il a donné lieu depuis à une énorme littérature et est devenu un objet mathématique à part entière: Il continue encore d'alimenter de nombreux travaux très fins, révélant petit à petit de nouvelles propriétés, et semble une source inépuisable de recherche.

Mais arrêtons-nous pour l'instant au début du siècle et formalisons schématiquement les propriétés alors connues du mouvement brownien. Le brownien est noté  $B(t, \omega)$ ,  $t$  désigne le temps et  $\omega$  modélise la dépendance par rapport au hasard, i.e  $\omega$  appartient à un espace de probabilité  $\Omega$  qu'il reste à construire. On notera simplement  $B_t$  quand il n'y aura pas de confusion possible et on adopte la convention  $B_0 = 0$ . Les propriétés alors connues ou soupçonnées sont:

- (1.1) A  $t$  fixé, pour tous  $s$ , la variable aléatoire  $B(t + s, \omega) - B(s, \omega)$  est une gaussienne centrée de variance  $\Delta t$ .
- (1.2) A  $\omega$  fixé, la fonction de  $t$  notée  $B_t$ , que l'on appelle une trajectoire du brownien, est continue.
- (1.3) Pour toute suite  $0 \leq t_1 < t_2 \dots < t_N$ , les v.a  $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_N} - B_{t_{N-1}})$  sont indépendantes.

Oublions les détails techniques pour l'instant (Construction de  $\Omega$ ...) et essayons de développer notre intuition de ce que peut être ce mouvement brownien, à l'aide d'un outil moderne et commode: l'ordinateur. Nous renvoyons à l'annexe I pour les détails sur la simulation numérique d'une gaussienne et oublions aussi pour le moment le problème délicat de simuler le hasard avec une machine déterministe. Pour visualiser une trajectoire du brownien sur  $[0, 1]$ , nous allons subdiviser cet intervalle de temps en  $0 = t_0 < t_1 < \dots < t_N = 1$  avec  $t_k = k/N$ , puis simuler les  $N$  v.a gaussiennes de variance  $1/N$ ,  $(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_N} - B_{t_{N-1}})$  d'après (1.1) et (1.3). Par commodité nous relierons ces  $N$  points par des segments de droite. Il faut donc imaginer une particule fictive qui suit un mouvement brownien (en  $x$ ) sur un intervalle de temps (en  $t$ ) de  $[0, 1]$ . Sur les deux figures suivantes apparaissent les courbes obtenues avec  $N = 100$  et  $N = 1000$ :



Le mouvement semble extrêmement chaotique, toutefois les deux points  $B_1$  ne sont curieusement pas superposables. Les courbes précédentes représentent donc deux trajectoires distinctes du brownien! En effet ces deux points ont été simulés à partir d'une suite pseudo-aléatoire *fixée* ( $U_1, U_2, \dots, U_n, \dots$ ) mais dans le premier cas  $B_1$  a été calculé à partir de  $U_{99}$  et  $U_{100}$  alors que dans le deuxième cas  $B_1$  a été calculé à partir de  $U_{999}$  et  $U_{1000}$  (voir Annexe I).

Nous voudrions observer la *même* trajectoire avec des pas de temps de plus en plus petits. Pour

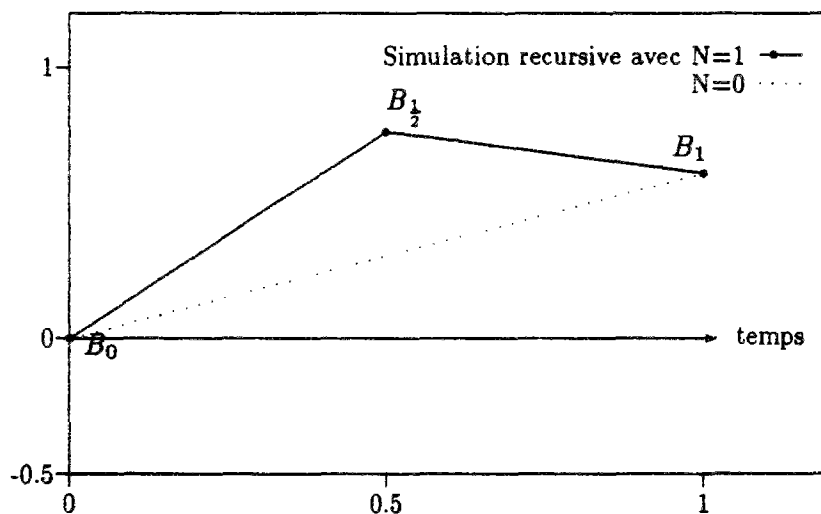
cela nous aimerions ne pas toucher aux points déjà connus et simuler de nouveaux points de la courbe pour *raffiner* la trajectoire approximante du brownien. On peut montrer facilement à partir de (1.1) à (1.3) que pour  $s < t$ ,

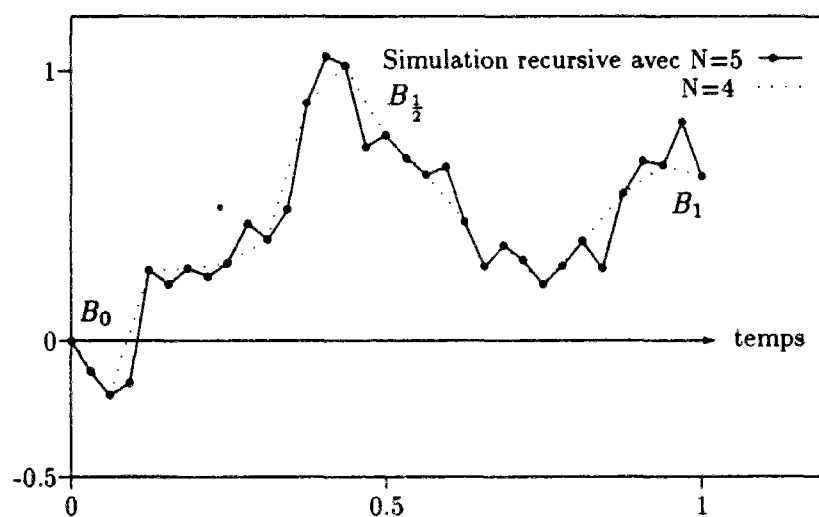
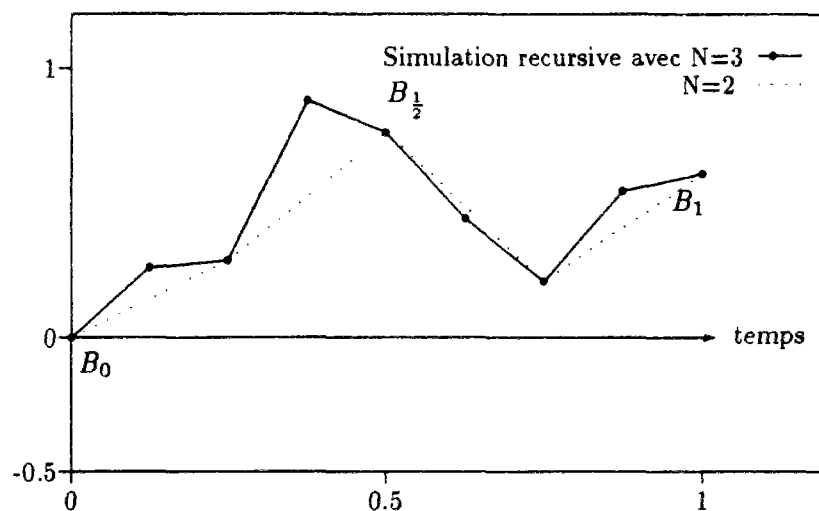
$$\mathcal{L}(B_{\frac{t+s}{2}} | B_s = x, B_t = y) = \mathcal{N}\left(\frac{x+y}{2}, \frac{t-s}{4}\right),$$

c'est à dire que, conditionnellement aux valeurs déjà prises par le brownien aux temps  $s$  et  $t$ , la loi du brownien au temps  $(t+s)/2$  est une gaussienne de moyenne  $(x+y)/2$  et de variance  $(t-s)/4$ .

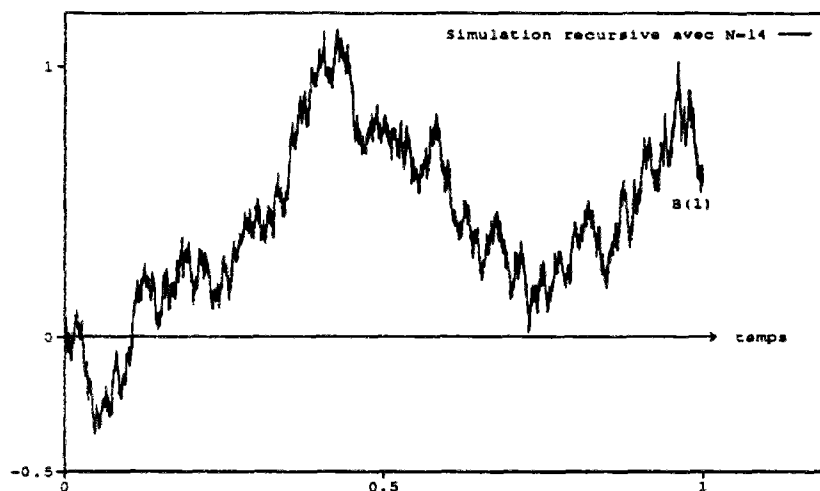
On va ainsi commencer par tirer les points  $B_0 = 0$ ,  $B_1$  de loi  $N(0,1)$ , puis simuler  $B_{1/2}$  par la propriété précédente. On fera alors de même pour obtenir  $B_{1/4}$  et  $B_{3/4}$  ... On dispose ainsi d'une *procédure récursive* qui permet, à la profondeur  $N$ , d'obtenir les valeurs de  $(B_{\frac{k}{2^N}})_{k=0 \dots 2^N}$ .

Les figures suivantes montrent la trajectoire obtenue pour  $N = 1$ ,  $N = 3$  et  $N = 5$ . En pointillé apparaît la trajectoire de l'étape précédente, i.e pour  $N = 0$ ,  $N = 2$  et  $N = 4$  respectivement.





Enfin sur la figure suivante est représentée la même trajectoire avec  $N = 14$ .



Que constate-t-on ? La nature extrêmement chaotique, fractale, du mouvement de la particule. Mais quand  $N$  tend vers l'infini, l'objet obtenu à la limite a-t-il de quelconques propriétés mathématiques ? Comment les découvrir ? Comment même formaliser correctement ce mouvement aléatoire ? Des questions similaires ont certainement été posées en ce tout début de siècle... et il a fallu attendre plusieurs dizaines d'années, Wiener en 1923 puis Paul Lévy en 1948 et bien d'autres encore, pour qu'elles trouvent des réponses satisfaisantes.

On sait maintenant très bien décrire ce comportement limite dans un formalisme adéquat. Donnons quelques propriétés classiques du mouvement brownien que nous empruntons à l'ouvrage de Karatzas-Shreve[88] et à qui nous renvoyons pour de plus amples renseignements.

Il existe un espace de probabilité  $\Omega$  sur lequel on peut définir un processus, appelé mouvement brownien, pour  $t \in [0, +\infty[$ , vérifiant les propriétés (1.1), (1.2), (1.3), et ces propriétés caractérisent ce processus en imposant  $B_0 = 0$ . De (1.1) à (1.3) on peut déduire une loi des grands nombres pour le mouvement brownien :

$$\lim_{t \rightarrow \infty} \frac{B_t}{t} = 0 \text{ p.s.}$$

On peut en fait préciser d'avantage le comportement à l'infini par la loi du logarithme itéré de Hincin (1933) qui peut se démontrer de façon élémentaire à partir de (1.1) et (1.3):

$$\limsup_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log \log t}} = 1 \text{ p.s.},$$

$$\liminf_{t \rightarrow \infty} \frac{B_t}{\sqrt{2t \log \log t}} = -1 \text{ p.s.}$$

Ces propriétés permettent aussi d'étudier le comportement local du brownien, par exemple en 0. Pour cela on montre que le processus défini par

$$X_0 = 0, X_t = tB_{\frac{1}{t}} \quad \forall t > 0$$

est un mouvement brownien car il vérifie (1.1) à (1.3). On en déduit alors que

$$\limsup_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \log \log \frac{1}{t}}} = 1 \text{ p.s.},$$

$$\liminf_{t \rightarrow 0} \frac{B_t}{\sqrt{2t \log \log \frac{1}{t}}} = -1 \text{ p.s.}$$

En particulier, partant de  $B_0 = 0$ , le brownien s'annule avec probabilité 1, une infinité de fois sur l'intervalle  $[0, \epsilon[$  pour tous  $\epsilon > 0$ ! Cette propriété peut se comprendre de façon intuitive si l'on pense à la simulation récursive vue précédemment. De plus on peut caractériser précisément sa régularité. Introduisons pour cela le module de continuité de Lévy (1937) pour un réel  $\delta > 0$ :

$$g(\delta) = \sqrt{2\delta \log \frac{1}{\delta}}.$$

On a alors :

$$\limsup_{\delta \rightarrow 0} \frac{1}{g(\delta)} \max_{0 < t-s \leq \delta} |B_t - B_s| = 1 \text{ p.s.}$$

En particulier, avec probabilité 1, les trajectoires du brownien ne sont nulles part monotones, ni dérivables, mais seulement holdériennes d'ordre  $1/2 - \epsilon$ . De plus, elles ne sont pas non plus à variation bornée mais à variation quadratique finie. Par exemple dans le cas d'une subdivision régulière  $t_0 = 0 < t_1 < \dots < t_N = 1$  de  $[0,1]$  de pas  $t_{k+1} - t_k = 1/N$ , on a

$$\lim_{N \rightarrow \infty} \sum_{k=0}^N |B_{t_{k+1}} - B_{t_k}| = +\infty \text{ p.s.},$$

$$\lim_{N \rightarrow \infty} \sum_{k=0}^N |B_{t_{k+1}} - B_{t_k}|^2 = 1 \text{ p.s.}$$

En particulier la longueur de la courbe brownienne sur  $[0,1]$  est infinie!

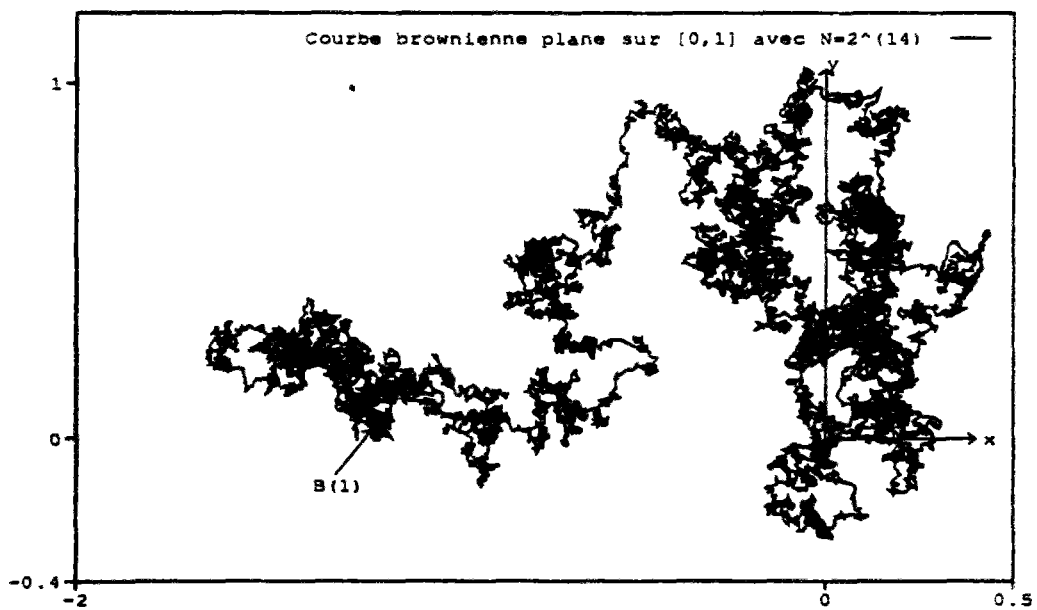
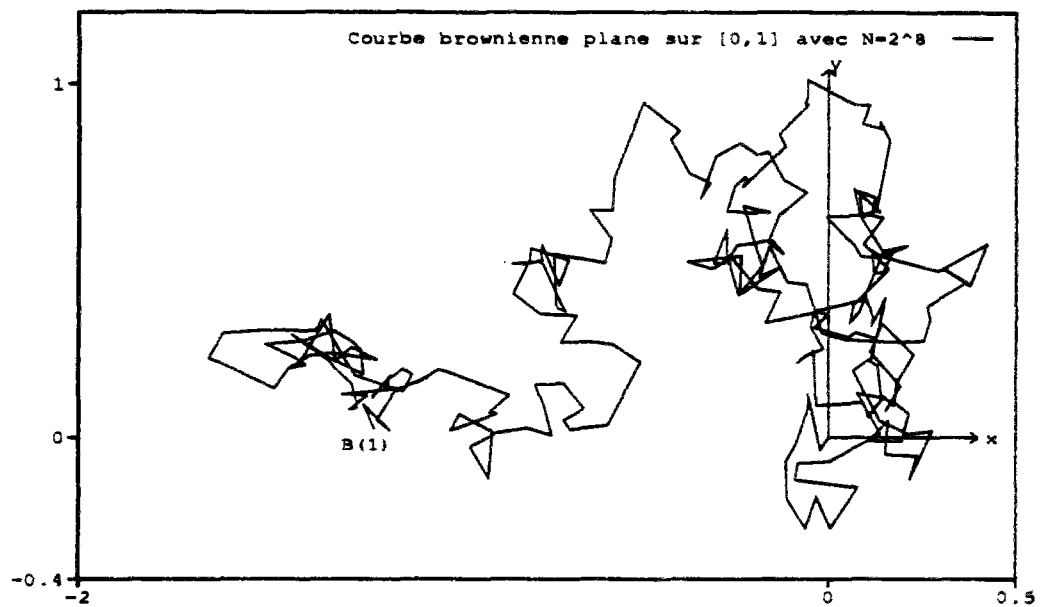
Il faut donc se garder de penser que les courbes précédentes représentent des trajectoires browniennes puisqu'elles ne rendent pas du tout compte des propriétés précédentes.

**Mouvement brownien dans  $\mathbf{R}^d$ .** Par définition  $B_t = (B_t^{(1)}, B_t^{(2)}, \dots, B_t^{(d)})$  où les  $B_t^{(i)}$  sont  $d$  browniens indépendants, est un mouvement brownien  $d$ -dimensionnel.

Reprenons notre outil moderne pour *stimuler notre intuition* de ce que peut être un brownien dans  $\mathbf{R}^d$ . Représentons dans le plan un brownien bi-dimensionnel en simulant deux discrétisations récursives de browniens indépendants comme précédemment. On imagine donc une particule de coordonnées  $(x = B_t^{(1)}, y = B_t^{(2)})$  qui suit un mouvement brownien plan au cours du temps  $t \in [0,1]$ .

Pour un pas de discrétisation de  $1/2^8$  on obtient une première représentation que l'on raffine ensuite en prenant un pas de  $1/2^{14}$ .





Que voit-on ? Un mouvement similaire à la dimension 1, c'est à dire très chaotique.

En fait il n'est pas similaire mais plus complexe. On peut montrer qu'en dimension 2 le mouvement brownien est récurrent, i.e il revient une infinité de fois visiter tous cercles aussi petits soient-ils. De plus il fait une infinité de tours dans un sens, puis dans l'autre, autour de certains points!

En dimension 3 et plus, le mouvement brownien est transcient, i.e il converge en norme vers l'infini quand le temps  $t$  tend vers l'infini, et ne vérifie donc pas la propriété de récurrence de la dimension 2. Il existe bien d'autres propriétés de ce type et on renvoie à l'excellente introduction au mouvement brownien de Le Gall[90] pour une présentation simple de ces propriétés.

On renvoie à l'ouvrage de Knight[81] pour une présentation assez complète avec des preuves élémentaires des résultats présentés jusqu'à maintenant, et à Karatzas-Shreve[88] pour un traitement moderne du mouvement brownien, ainsi qu'aux références incluses dans ces ouvrages. Nous précisons maintenant quelques notations et donnons quelques commentaires.

**Commentaires et notations.** La simulation itérative, que nous avons utilisé pour avoir une représentation "naïve" du brownien, se formalise et permet de donner une construction complète du brownien (voir Karatzas-Shreve[88] 2.4.D). La méthode de simulation récursive vue précédemment est due à Lévy(1948) et Ciesielski(1961) et permet une autre construction indépendante du mouvement brownien (voir Karatzas-Shreve[88] 2.3.).

Dans toute la suite on supposera que l'on dispose d'un espace de probabilité  $(\Omega, \mathcal{A}, \mathbf{P})$  sur lequel est défini une version régulière (continue) du mouvement brownien notée  $B_t$ . On renvoie par exemple à Karatzas-Shreve[88] pour la construction d'un tel espace. Par convention on prendra  $B_0 = 0$  et  $\mathbf{E}^x(f(B_t))$  représentera  $\mathbf{E}(f(B_t + x))$ .

Par la suite on notera

$$\mathcal{F}_t = \sigma(B_s, s \leq t)$$

la  $\sigma$ -algèbre engendrée par le brownien jusqu'en  $t$ , qui représente intuitivement l'information disponible sur la trajectoire du brownien jusqu'au temps  $t$ .

Rappelons aussi la définition des temps d'arrêts, qui jouent un rôle majeur dans la théorie des processus stochastiques. On dit que  $\tau$  est un temps d'arrêt par rapport à la filtration  $\mathcal{F}_t$  si

$$\forall t \geq 0, \{\tau \leq t\} \in \mathcal{F}_t.$$

Un exemple typique est le temps d'atteinte d'un niveau  $a > 0$ :

$$\tau_a(\omega) = \inf\{t \geq 0, B_t(\omega) = a\}.$$

Un temps d'arrêt est donc intuitivement une information sur le brownien qui *n'anticipe pas* sur le futur.

Rappelons enfin quelques propriétés qui jouent un rôle essentiel dans l'étude du mouvement brownien, à savoir

- La propriété de changement d'échelle (ou scaling):  $\forall c > 0$ ,  $\frac{1}{c}B_{c^2t}$  est un mouvement brownien. Ce qui veut dire qu'une observation à la loupe donne le même "motif aléatoire" à une échelle de temps près.

- Le brownien  $B_t$  est une martingale, i.e  $B_t$  est intégrable et, pour  $s < t$ ,

$$E(B_t | \mathcal{F}_s) = B_s.$$

- Le brownien vérifie la propriété de Markov forte, i.e pour tous temps d'arrêt  $\tau$  p.s fini, le processus défini par

$$W_t = B_{t+\tau} - B_\tau$$

est un mouvement brownien. Ainsi, arrêté à n'importe quel instant (raisonnable), le brownien repart "sans mémoire" de son passé avant cet instant.

Revenons maintenant à nos simulations. Nous avons vu une méthode de **simulation itérative** du brownien très simple s'inspirant de (1.1) et (1.3). Si cette méthode est très classique, on peut toutefois constater qu'elle rend très mal compte des propriétés fines du brownien. Nous avons aussi vu une méthode de **simulation récursive** qui contient *en puissance* ces propriétés fines et permet de raffiner une trajectoire en construisant ce que l'on appelle des ponts browniens.

Toutefois, si l'on veut observer le comportement très fin du brownien sur tout l'intervalle  $[0, 1]$ , la méthode précédente devient un peu fastidieuse par le nombre très élevé de points de discrétisation. De plus elle ne donne pas de renseignement sur le comportement du brownien *entre* ces points de discrétisation et ne rend donc pas compte, par exemple, de la loi du logarithme itéré locale vue précédemment.

Il existe d'autres simulations possibles du brownien, par exemple comme limite de marches aléatoires (Karatzas-Shreve[88] par exemple) mais elles présentent aussi en général l'inconvénient de mal rendre compte du *comportement spatial* du brownien entre les points de discrétisation.

A l'inverse nous allons voir au chapitre suivant une autre simulation du brownien moins classique: Elle consiste à suivre "précisément" dans l'espace une trajectoire brownienne et permet nous le verrons de mieux rendre compte de sa structure fine.

## Chapitre 2

# Simulation spatialement contrôlée du brownien

## 2.1 Introduction.

Soit  $(B_t)$  un brownien  $d$ -dimensionnel issu de zéro. On note  $|\cdot|$  la norme du max définie pour tous  $x = (x^1, \dots, x^d)$  de  $\mathbf{R}^d$  par

$$|x| = \max_i |x^i|.$$

On définit, pour  $a > 0$ , le temps d'arrêt

$$\tau_a = \inf\{t > 0, |B_t| = a\},$$

qui représente donc le premier temps de sortie du cube  $[-a, a]$ .

Le but de ce chapitre 2 est de proposer une méthode de simulation aussi exacte que possible du couple  $(\tau_a, B_{\tau_a})$ .

Quelles sont les motivations d'une telle étude ?

C'est de permettre une simulation du brownien qui soit *spatialement contrôlée*. Dans certains cas comme l'étude du comportement local du brownien, l'approximation trajectorielle fine du temps de sortie d'un ouvert..., la simulation itérative classique comme suite d'accroissements indépendants de gaussiennes à des temps discrétisés est insuffisante. En effet on ne dispose d'aucune information sur le comportement du brownien entre les points de discrétisation alors que l'on sait, par exemple par la propriété du logarithme itéré, que le comportement du brownien entre ces points est extrêmement irrégulier (voir Chapitre 1).

Nous commençons par aborder le cas de la dimension 1 en 2.2, puis la dimension  $d$  en 2.3. Nous traiterons ensuite le cas d'un brownien avec drift en 2.4.

Enfin nous montrerons en 2.5 comment utiliser ces résultats pour simuler le brownien à sa sortie d'un ouvert borné quelconque, et nous ferons une application numérique dans le cas du cercle.

Ces résultats seront aussi utilisés dans le chapitre 6 sur l'approximation spatiale des diffusions.

## 2.2 Etude du temps de sortie $\tau_a$ d'un cube, dimension 1.

### 2.2.1 Quelques propriétés de $\tau_a$ .

On se place pour le moment en dimension 1, la loi de  $\tau_a$  est alors bien connue. Rappelons-en quelques propriétés classiques qui nous seront utiles ultérieurement. Nous avons la

**Proposition 1.**

$$(i) \tau_a = a^2 \tau_1 \text{ en loi.}$$

$$(ii) \forall x > 0, \mathbf{E} \exp(-x\tau_a) = \frac{1}{\cosh(a\sqrt{2x})}.$$

$$(iii) \forall x \in [0, \frac{\pi^2}{8a^2}], \mathbf{E} \exp(x\tau_a) = \frac{1}{\cos(a\sqrt{2x})}.$$

$$(iv) \forall p \geq 0, \exists C_p < +\infty, \mathbf{E} \tau_a^p = C_p a^{2p}.$$

**Preuve:**

(i) Par changement d'échelle on obtient les égalités suivantes en loi

$$\tau_a = \inf\{t > 0, |B_{\frac{t}{a^2}}| = 1\} = a^2 \inf\{t > 0, |B_t| = 1\} = a^2 \tau_1.$$

Il suffit donc de prouver les trois propriétés suivantes pour  $a = 1$ . Les formules (ii) et (iii) proviennent de Itô-Mckean[65]1.7 pb 6. Pour prouver (iv) on remarque que, d'après (iii),  $\tau_a$  admet des moments exponentiels donc  $\exists C_p < +\infty$ ,  $\mathbf{E} \tau_1^p = C_p$ .

□

On peut remarquer que la loi du temps d'atteinte

$$T_a = \inf\{t > 0, B_t = a\}$$

est bien connue et penser utiliser la propriété  $\tau_a = T_a \wedge T_{-a}$  pour simuler la variable aléatoire  $\tau_a$ . Malheureusement  $T_a$  et  $T_{-a}$  ne sont pas indépendantes. Aussi, pour simuler la v.a.  $\tau_a$ , on utilise la propriété suivante qui donne la densité du temps de sortie. Une méthode de rejet permet alors de faire la simulation comme nous le verrons par la suite. Nous avons la

**Proposition 2.** La densité  $f_a$  de  $\tau_a$  est donnée par les deux formules suivantes:

$$f_a(t) = \frac{\pi}{2a^2} \sum_{n=0}^{\infty} (-1)^n (2n+1) \exp\left(-\frac{(2n+1)^2 \pi^2 t}{8a^2}\right), \quad t > 0. \quad (2.1)$$

$$f_a(t) = \sqrt{\frac{2}{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{(2n+1)a}{t^{3/2}} \exp\left(-\frac{a^2(2n+1)^2}{2t}\right), \quad t > 0. \quad (2.2)$$

**Preuve:** La première formule est classique, de Feller[71] X.5 (5.9) on a

$$\mathbf{P}(\tau_a > t) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp\left(-\frac{(2n+1)^2 \pi^2 t}{8a^2}\right).$$

Par dérivation on obtient (2.1). Puis de Feller[71] X.5 (5.8) on obtient

$$\mathbf{P}(\tau_a > t) = 2 \sum_{-\infty}^{+\infty} \mathbf{N}\left(\frac{(4k+1)a}{\sqrt{t}}\right) - \mathbf{N}\left(\frac{(4k-1)a}{\sqrt{t}}\right)$$

où  $\mathbf{N}(x)$  désigne la fonction de répartition d'une loi normale centrée réduite. Par dérivation il vient

$$f_a(t) = \frac{a}{\sqrt{2\pi t^3}} \sum_{-\infty}^{+\infty} (g(4k+1) - g(4k-1))$$

où l'on a noté

$$g(x) = x \exp\left(-\frac{a^2 x^2}{2t}\right), \quad \forall x \in \mathbf{R}^d.$$

La fonction  $g$  étant impaire, on a facilement

$$\sum_{-\infty}^{+\infty} (g(4k+1) - g(4k-1)) = 2 \sum_0^{+\infty} (-1)^k g(2k+1),$$

qui donne le résultat. Cette formule peut aussi s'obtenir à partir de l'exercice 8.11 de Karatzas-Shreve[88] p99.

□

On peut aussi passer de (2.1) à (2.2) en utilisant la formule de sommation de Poisson:

$$\sum_{-\infty}^{+\infty} \exp(-\pi(k+z)^2 a) = \frac{1}{\sqrt{a}} \sum_{-\infty}^{+\infty} \exp\left(-\frac{\pi k^2}{a} + 2i\pi k z\right).$$

**Remarque 1** Pour la simulation, nous verrons que les deux formules de la densité sont nécessaires car la première série converge bien pour  $t$  grand et la deuxième converge bien pour  $t$  petit.

### 2.2.2 Simulation de $(\tau_a, B_{\tau_a})$ .

- Pour simuler  $B_{\tau_a}$ , il suffit de remarquer que  $B_{\tau_a}$  est indépendant de  $\tau_a$  et que sa loi est  $\frac{1}{2}\delta_a + \frac{1}{2}\delta_{-a}$ .
- Pour simuler  $\tau_a$ , la densité étant donnée par une série convergente, nous utilisons une méthode de rejet. Une telle méthode est exposée dans Devroye[86] IV.5 p 151. Rappelons le principe de simulation (Nous corrigeons une erreur apparaissant dans notre édition de cet ouvrage):

Supposons que la densité à simuler  $f(x)$  s'écrive sous la forme

$$f(x) = \sum_{n=0}^{\infty} s_n(x) \leq c h(x),$$

où  $h$  est une densité de probabilité facilement simulable,  $c \geq 1$ , avec de plus

$$\left| \sum_{k=n}^{\infty} s_k(x) \right| \leq R_n(x),$$

où  $R_n(x)$  converge vers 0 quand  $n$  tend vers l'infini.

L'idée est que l'on peut adapter une méthode de rejet classique car  $f$  peut être approchée avec autant de précision souhaitée. On simule la v.a  $X$  de densité  $f(x)$  par l'algorithme:

*repeat*

Générer  $X$  de densité  $h$ , puis  $U$  uniforme sur  $[0, 1]$ .

$W := Uch(X)$

$s := 0$  ,  $n := -1$

```

repeat
     $n := n + 1, s := s + s_n(X)$ 
until  $|S - W| > R_{n+1}(X)$ 
until  $W \leq S$ 
return  $X$ .

```

$X$  est alors une réalisation de la v.a  $\tau_a$ .

C'est cette méthode que nous allons employer en scindant la densité en deux parties pour des raisons de convergence. Pour cela on utilise la

**Définition 1** Si  $F_a(t)$  désigne la fonction de répartition de  $\tau_a$  et  $f_a(t)$  sa densité, on définit les densités  $p_1(t)$  et  $p_2(t)$  par

$$p_1(t) = \frac{1}{F_a(a^2)} f_a(t) \mathbf{1}_{[0, a^2]}(t)$$

et

$$p_2(t) = \frac{1}{1 - F_a(a^2)} f_a(t) \mathbf{1}_{[a^2, \infty]}(t) .$$

Montrons comment on peut simuler les densités  $p_1(t)$  et  $p_2(t)$ . Pour  $p_1(t)$ , on utilise la propriété suivante.

**Proposition 3.** La densité  $p_1$  s'écrit

$$p_1(t) = \sum_{k=0}^{\infty} s_k(t) \leq C_1 h_1(t)$$

avec

$$|\sum_{k=n}^{\infty} s_k(t)| \leq R_n(t).$$

Les fonctions  $s_n(t)$ ,  $C_1$ ,  $h_1(t)$ ,  $R_n(t)$  sont données par

$$s_n(t) = (-1)^n \sqrt{\frac{2}{\pi}} \frac{(2n+1)a}{t^{3/2} F_a(a^2)} \exp\left(-\frac{a^2(2n+1)^2}{2t}\right), \quad t \in [0, a^2],$$

$$C_1 = \frac{4\sqrt{2e}}{(e-1)^2 F_a(a^2) \sqrt{\pi}},$$

$$h_1(t) = \frac{a^2}{2t^2} \exp\left(\frac{1}{2}\left(1 - \frac{a^2}{t}\right)\right),$$

$$R_n(t) = |s_n(t)|.$$

De plus, si  $U$  est uniformément répartie sur  $[0, 1]$ , alors la v.a

$$X_1 = \frac{a^2}{1 - 2\log(U)}$$

suit une loi de densité  $h_1$ .



**Preuve:** On a pour  $t \in [0, a^2]$ ,

$$s_n(t) = h_1(t) \frac{1}{F_a(a^2)} \frac{\sqrt{2}}{\sqrt{e\pi}} \frac{2t^2}{a^2} (-1)^n \frac{(2n+1)a}{t^{3/2}} \exp\left(-\frac{a^2 2n(n+1)}{t}\right)$$

donc

$$p_1(t) \leq h_1(t) \sum_0^\infty \frac{2\sqrt{2}}{F_a(a^2)\sqrt{e\pi}} (2n+1) \exp(-2n(n+1)).$$

Or

$$\sum_0^\infty (2n+1) \exp(-2n(n+1)) \leq 2 \sum_0^\infty (n+1) \exp(-(n+1)) = \frac{2e}{(e-1)^2}$$

donc

$$s_n(t) \leq C_1 h_1(t).$$

On vérifie que  $h_1$  est bien une densité. En inversant sa fonction de répartition  $F_1(x)$  on obtient

$$F_1^{-1}(x) = \frac{a^2}{1 - 2\log(x)}$$

qui permet de simuler, par la méthode d'inversion de la fonction de répartition (Devroye[86]), la v.a  $X_1$  de densité  $h_1(x)$ . Enfin la série  $s_n(x)$  est alternée. En effet si on note

$$f(x) = x \exp(-mx^2) \text{ pour } x \in \mathbf{R}^d,$$

alors la fonction  $f$  est décroissante pour  $x > (2m)^{-1/2}$ . Comme ici  $m = \frac{a^2}{2t}$ , on a  $(2m)^{-1/2} \leq 1$  et donc  $f(2n+1) \geq f(2n+3)$ . On obtient ainsi la formule donnant  $R_n(t)$ .

□

On procède ensuite de manière analogue pour la densité  $p_2(t)$  sur  $[a^2, +\infty[$ . On renvoie à Devroye[86] IV.5 p 151 pour l'algorithme de simulation quand la densité se factorise avec une série alternée de premier terme 1 (qui est une simple variante de l'algorithme précédent, mais une erreur apparait dans l'édition que nous possédons: il faut lire 1 à la place de  $c$ ). Nous avons la

**Proposition 4.** *On peut écrire la densité  $p_2(t)$  sous la forme*

$$p_2(t) = C_2 h_2(t) \sum_{n=0}^\infty (-1)^n s_n(t), \quad \forall t \in [a^2, +\infty[.$$

avec

$$s_n(t) = (2n+1) \exp\left(-\frac{n(n+1)t\pi^2}{2a^2}\right),$$

$$C_2 = \frac{4 \exp(-\pi^2/8)}{\pi(1 - F_a(a^2))}.$$

$$h_2(t) = \frac{\pi^2}{8a^2} \exp\left(\frac{\pi^2}{8}\left(1 - \frac{t}{a^2}\right)\right), \quad \forall t \in [a^2, +\infty[.$$

De plus, si  $U$  est uniformément répartie sur  $[0, 1]$  alors la v.a

$$X_2 = a^2\left(1 - \frac{8}{\pi^2} \log(U)\right)$$

suit une loi de densité  $h_2$ .

**Preuve:** D'après la proposition 2 on a

$$p_2(t) = \frac{1}{1 - F_a(a^2)} \frac{\pi}{2a^2} \sum_{n=0}^{\infty} (-1)^n (2n+1) \exp\left(-\frac{(2n+1)^2 \pi^2 t}{8a^2}\right)$$

et donc

$$p_2(t) = C_2 h_2(t) \sum_{n=0}^{+\infty} s_n(t).$$

On vérifie que  $s_n$  est bien décroissante comme à la proposition 3, que  $h_2(t)$  est bien une densité de probabilité et que sa fonction de répartition  $F_2(x)$  vérifie

$$F_2^{-1}(x) = a^2\left(1 - \frac{8}{\pi^2} \log(1-x)\right).$$

On peut donc simuler  $X_2$  par inversion de la fonction de répartition.

□

**Simulation de  $\tau_a$ :** Pour simuler  $\tau_a$  à partir des deux densités précédentes on procède de la façon suivante:

On génère  $U$  uniforme sur  $[0, 1]$ . Soit  $p = F_a(a^2)$ . Deux cas se présentent:

- Si  $U < p$ , on simule  $X$  de densité  $p_1$ .
- Si  $U \geq p$ , on simule  $X$  de densité  $p_2$ .

$X$  est alors une réalisation de la v.a  $\tau_a$ . Enfin il reste à estimer  $p = F_a(a^2)$ . Par intégration on obtient

$$F_a(a^2) = 1 - \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} \exp\left(-\frac{(2n+1)^2 \pi^2}{8}\right).$$

Pour les applications cette série alternée, qui converge très rapidement, est tronquée quand son reste est inférieur à  $10^{-10}$ .

**Validation expérimentale.** Pour vérifier expérimentalement l'algorithme de simulation de  $\tau_a$ , nous aurions pu utiliser (ii) de la Proposition 1. Nous utilisons toutefois une autre propriété qui a l'avantage de bien se généraliser aux dimensions supérieures et sera ainsi utilisée en 2.3 et 2.4. D'après le théorème d'arrêt, on peut remarquer que  $EB_{\tau_a}^2 = E\tau_a$ . On a donc la propriété

$$E(\tau_a - a^2) = 0.$$

C'est cette propriété que nous avons choisie de vérifier par Monte-Carlo. La courbe de la figure 2.1 illustre la convergence vers zéro de  $E(\tau_a - a^2)$  par Monte Carlo.

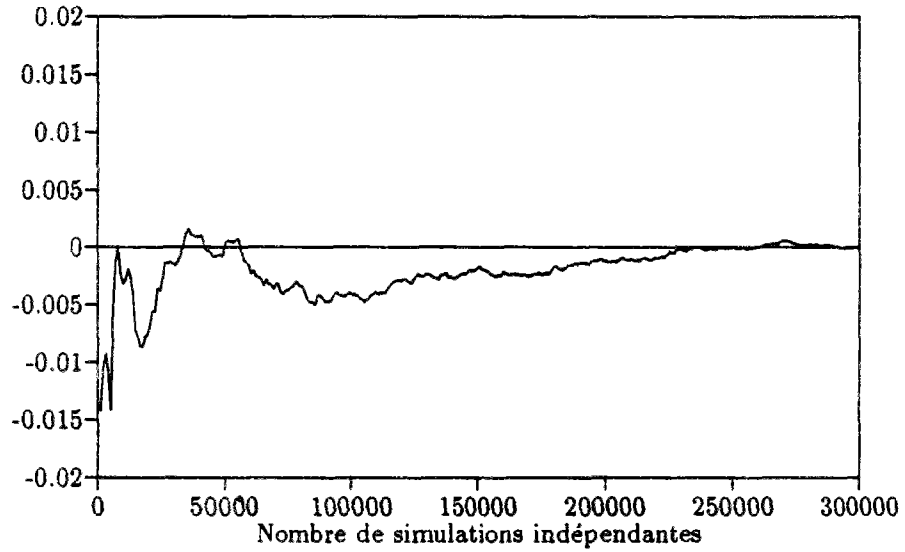


Figure 2.1: Calcul par Monte Carlo de  $E(\tau_1 - 1)$

**Performances de l'algorithme:** Le nombre d'itérations de la boucle de rejet dans l'algorithme de simulation est théoriquement  $pC_1 + (1 - p)C_2$  soit environ 2.15. De plus le temps de simulation d'une réalisation de  $\tau_a$  est de  $5.6 \cdot 10^{-4}$  cpu sur un Spark2, avec une variance très faible, et représente donc (annexe1) en moyenne environ 20 fois le temps de génération d'un nombre pseudo-aléatoire, ce qui reste négligeable en pratique.

Si toutefois une implémentation très rapide était nécessaire, un choix plus fin des constantes  $C_1$  et  $C_2$  ainsi que l'utilisation de méthodes mixtes exposées dans Devroye[86], pourrait diminuer très notablement les temps de simulation.

## 2.3 Etude du temps de sortie $\tau_a$ d'un cube, dimension $d$ .

Soit  $B_t = (B_t^1, \dots, B_t^d)$  un brownien  $d$ -dimensionnel issu de zéro. On note pour  $a > 0$ ,

$$\tau_a = \inf\{t > 0, |B_t| = a\} \text{ avec } |x| = \max_i |x^i| \text{ pour } x = (x^1, \dots, x^d),$$

et pour  $i \in [1, d]$ ,

$$\tau_a^i = \inf\{t > 0, |B_t^i| = a\}.$$

On a bien sûr

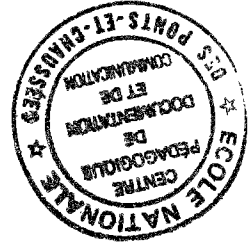
$$\tau_a = \min_{i=1..d} \tau_a^i.$$

Il est alors facile de déduire par changement d'échelle que

$$\forall p > 0, \exists C_{p,d} \leq 1, E(\tau_a)^p = C_{p,d} a^{2p}.$$

Des propriétés analogues à celle de la proposition 1 peuvent être obtenues.

Nous nous proposons maintenant de simuler le  $(d+1)$ -uplet  $(\tau_a, B_{\tau_a})$ .



### 2.3.1 Un algorithme de simulation de $(\tau_a, B_{\tau_a})$ .

- **Simulation de  $\tau_a$ .** Il suffit de simuler, par indépendance, les  $d$  variables aléatoires  $\tau_a^i$  et d'en prendre le minimum. On est alors ramené à  $d$  simulations en dimension 1 et la méthode exposée en 2.2 s'applique.
- **Simulation de  $B_{\tau_a}$ .** On vient de simuler  $\tau_a$ . P.p.s, si  $i_0$  est un indice tel que  $\tau_a = \tau_a^{i_0}$ , alors on a  $\forall j \neq i_0, \tau_a < \tau_a^j$ , par indépendance des composantes du brownien. Simulons donc aussi  $i_0$ , qui est indépendant de  $\tau_a$ , et équirépartis sur  $[1, d]$ .

Il est clair que  $B_{\tau_a}^{i_0} = \epsilon a$  où la variable aléatoire  $\epsilon$  a pour loi  $\frac{1}{2}\delta_1 + \frac{1}{2}\delta_{-1}$  et est indépendante de  $i_0$  et de  $\tau_a$ .

Pour la simulation des autres composantes du brownien il faut connaître la loi de  $B_{\tau_a}^j$  conditionnellement à  $\tau_a$  et à  $j \neq i_0$ . Comme nous allons le voir, il suffit pour cela de connaître la loi de  $(B_t^j | t < \tau_a^j)$ , rappelée en Proposition 5, puis montrer que l'on peut substituer dans cette formule  $t$  par  $\tau_a$ , c'est l'objet de la Proposition 6.

**Proposition 5.** Soit  $B_t$  un brownien réel issu de zéro. Si  $\tau_a = \inf\{t > 0, |B_t| = a\}$  alors pour  $x \in [-a, a]$ ,

$$P(B_t \in dx \mid t < \tau_a) = \frac{1}{a\nu_a(t)} \sum_{k=0}^{+\infty} \exp\left(-\frac{(2k+1)^2\pi^2 t}{8a^2}\right) \cos\left(\frac{(2k+1)\pi x}{2a}\right) dx,$$

avec  $\nu_a(t)$  donné par

$$\nu_a(t) = P(\tau_a > t) = \frac{4}{\pi} \sum_{k=0}^{+\infty} \frac{(-1)^k}{2k+1} \exp\left(-\frac{(2k+1)^2\pi^2 t}{8a^2}\right).$$

**Preuve:** Notons pour  $b \in \mathbb{R}$ ,

$$m_b = \inf\{t > 0, B_t = b\}.$$

De Itô-McKean[65] 1.7 Pb 8 on a

$$\mathbf{P}_a(B_t \in dx, t < m_0, t < m_1) = 2 \sum_{k=0}^{+\infty} \sin(k\pi a) \sin(k\pi x) \exp(-k^2 \pi^2 t/2) dx.$$

Par translation on obtient

$$\mathbf{P}_0(B_t \in dx, t < m_{-\frac{1}{2}}, t < m_{\frac{1}{2}}) = 2 \sum_{k=0}^{+\infty} \cos((2k+1)\pi x) \exp(-(2k+1)^2 \pi^2 t/2) dx.$$

On utilise alors un argument de scaling,

$$\begin{aligned} \mathbf{P}_0(B_t \in db, t < \tau_a) &= \mathbf{P}_0(B_t \in db, \sup_{s \leq t} |B_s| < a) \\ &= \mathbf{P}_0(B_{\frac{t}{4a^2}} \in d(\frac{b}{2a}), \sup_{s \leq t} |B_{\frac{s}{4a^2}}| < \frac{1}{2}) \\ &= \mathbf{P}_0(B_u \in d(\frac{b}{2a}), \sup_{v \leq u} |B_v| < \frac{1}{2}) \text{ avec } u = \frac{t}{4a^2}. \end{aligned}$$

On obtient ainsi la première formule. La deuxième formule a été vue dans la preuve de la proposition 2. Elle peut aussi s'obtenir en calculant

$$\int_{-a}^a \mathbf{P}(B_t \in dx, t < \tau_a) dx.$$

□

Notons qu'une autre formule peut être obtenue, voir Itô-McKean[65] 1.7 pb 8 ou Karatzas-Shreve[87] (chapitre 2 propriété 8.10).

Pour simuler  $B_{\tau_a}^j$  sachant  $\tau_a$  et  $\tau_a^j > \tau_a$  nous utilisons alors la

**Proposition 6.** Soit  $\Theta_f^j(t) = \mathbf{E}(f(B_t^j) | t < \tau_a^j)$  qui est donné par la proposition 5. On a pour tous  $j$

$$\mathbf{E}(f(B_{\tau_a}^j) | \mathbf{1}_{\{\tau_a < \tau_a^j\}}, \tau_a) = \Theta_f^j(\tau_a) \mathbf{1}_{\{\tau_a < \tau_a^j\}} + \frac{1}{2}(f(a) + f(-a)) \mathbf{1}_{\{\tau_a = \tau_a^j\}}.$$

**Preuve:** Donnons la démonstration dans le cas  $d = 3$ , la démonstration générale est analogue. D'après la remarque au début de 2.3.1 il suffit de montrer que, pour  $h$  bornée, on a

$$\mathbf{E}(f(B_{\tau_a}^j) \mathbf{1}_{\{\tau_a = \tau_a^j\}} \mathbf{1}_{\{j \neq i\}} h(\tau_a)) = \mathbf{E}(\Theta_f^j(\tau_a) \mathbf{1}_{\{\tau_a = \tau_a^j\}} \mathbf{1}_{\{j \neq i\}} h(\tau_a)).$$

Remarquons que, sauf sur un ensemble de mesure nulle, si  $i$  est l'indice aléatoire pour lequel  $\tau_a = \tau_a^i$  alors  $\forall j \neq i$ ,  $\tau_a^i < \tau_a^j$ . Dans ce qui suit on suppose par exemple que  $i = 2$ ,  $j = 1$ . En effet, par symétrie, les autres cas se traitent de la même façon.

Les égalités suivantes se justifient en remarquant que  $\tau_a^1$ ,  $\tau_a^2$  et  $\tau_a^3$  sont indépendantes deux à deux.. On a

$$\begin{aligned} \mathbf{E}(f(B_{\tau_a}^1)1_{\{\tau_a=\tau_a^2\}}h(\tau_a)) &= \mathbf{E}(f(B_{\tau_a}^1)1_{\{\tau_a^1>\tau_a^2\}}1_{\{\tau_a^3>\tau_a^2\}}h(\tau_a^2)) \\ &= \mathbf{E}(\mathbf{E}(f(B_{\tau_a}^1)1_{\{\tau_a^1>\tau_a^2\}}|\tau_a^2, \tau_a^3) 1_{\{\tau_a^3>\tau_a^2\}}h(\tau_a)) \\ &= \mathbf{E}(\mathbf{E}(f(B_{\tau_a}^1)1_{\{\tau_a^1>\tau_a^2\}}|\tau_a^2) 1_{\{\tau_a^3>\tau_a^2\}}h(\tau_a)) \end{aligned}$$

De plus si on note

$$\Psi(t) = \mathbf{E}(f(B_t^1)1_{\{\tau_a^1>t\}})$$

alors on a

$$\begin{aligned} \mathbf{E}(f(B_{\tau_a}^1)1_{\{\tau_a=\tau_a^2\}}h(\tau_a^2)) &= \mathbf{E}(\Psi(\tau_a^2)1_{\{\tau_a^3>\tau_a^2\}}h(\tau_a^2)) \\ &= \mathbf{E}(\Theta_f^1(\tau_a^2)1_{\{\tau_a^3>\tau_a^2\}}\mathbf{E}(1_{\{\tau_a^1>\tau_a^2\}}|\tau_a^2) h(\tau_a^2)). \end{aligned}$$

Puis en déconditionnant il vient

$$\mathbf{E}(f(B_{\tau_a}^1)1_{\{\tau_a=\tau_a^2\}}h(\tau_a^2)) = \mathbf{E}(\Theta_f^1(\tau_a^2)1_{\{\tau_a=\tau_a^2\}}h(\tau_a^2)).$$

□

Etudions maintenant la simulation numérique de  $B_{\tau_a}^j$ .

**Un algorithme de simulation de  $B_{\tau_a}$ .** On a vu comment simuler  $\tau_a$  précédemment.

Pour simuler  $B_{\tau_a}$ , d'après les deux propositions précédentes, il suffit de savoir simuler la densité de la proposition 5. Cette densité est donnée par une série convergente, on procède donc conformément à l'algorithme rappelé en 2.2 Plus précisément on a la

**Proposition 7.** Si on note  $q(x)$  la densité définie par

$$q(x) = \frac{2}{a\nu_a(t)} \sum_{k=0}^{+\infty} \exp(-\frac{(2k+1)^2\pi^2 t}{8a^2}) \cos(\frac{(2k+1)\pi x}{2a}) \text{ pour } x \in [0, a]$$

alors on a

$$q(x) = \sum_{k=0}^{\infty} s_n(x) \leq Ch(x)$$

avec

$$| \sum_{k=n}^{\infty} s_k(x) | \leq R_n(x).$$

Les fonctions  $s_n(x)$ ,  $C$ ,  $h(x)$ ,  $R_n(x)$  sont données par

$$\forall x \in [0, a] \quad s_n(x) = \frac{2}{a\nu_a(t)} (-1)^n \exp\left(-\frac{(2n+1)^2 \pi^2 t}{8a^2}\right) \cos\left(\frac{(2n+1)\pi x}{2a}\right), \quad x \in [0, a]$$

$$C = \frac{8}{\pi \nu_a(t) (1 - \exp(-t\pi^2/(2a^2)))},$$

$$h(x) = \frac{\pi}{4a} \sqrt{2} \cos\left(\frac{\pi x}{2a} - \frac{\pi}{4}\right),$$

$$R_n(x) = \frac{2}{a\nu_a(t)} \exp\left(-\frac{nt\pi^2}{2a^2}\right) \frac{1}{1 - \exp(-t\pi^2/(2a^2))}.$$

De plus, si  $U$  est uniformément répartie sur  $[0, 1]$  alors la v.a  $X$  suit une loi de densité  $h$  en prenant

$$X = \frac{a}{2} + \frac{2a}{\pi} \operatorname{Arcsin}\left(\frac{2U - 1}{\sqrt{2}}\right).$$

**Preuve:** On a

$$q(x) \leq \frac{2}{a\nu_a(t)} \left( \cos\left(\frac{\pi x}{2a}\right) + \sin\left(\frac{\pi x}{2a}\right) \right) \sum_{k=0}^{+\infty} \exp\left(-\frac{(2k+1)^2 \pi^2 t}{8a^2}\right),$$

puis

$$q(x) \leq \frac{2}{a\nu_a(t)} \sqrt{2} \cos\left(\frac{\pi x}{2a} - \frac{\pi}{4}\right) \sum_{k=0}^{+\infty} \exp\left(-\frac{4k\pi^2 t}{8a^2}\right)$$

et donc

$$q(x) \leq Ch(x).$$

On procède de même pour obtenir  $R_n(x)$ .

Enfin on vérifie que  $h(x)$  est bien une densité et que sa fonction de répartition  $H(x)$  vérifie

$$H^{-1}(x) = \frac{a}{2} + \frac{2a}{\pi} \operatorname{Arcsin}\left(\frac{2x - 1}{\sqrt{2}}\right).$$

**Remarque 2** En pratique le calcul de  $\nu_a(t)$  défini à la proposition 5 est effectué avec une précision de  $10^{-10}$  (il est facile d'avoir un majorant simple du reste de la série).

**Validation expérimentale et performances.** Les algorithmes permettant de simuler le couple  $(\tau_a, B_{\tau_a})$  étant assez complexes, il est nécessaire de faire quelques vérifications expérimentales pour s'assurer de la justesse de leur implémentation informatique.

On peut remarquer qu'en appliquant le théorème d'arrêt on a

$$\mathbb{E}(|B_{\tau_a}|^2 - d \tau_a) = 0.$$

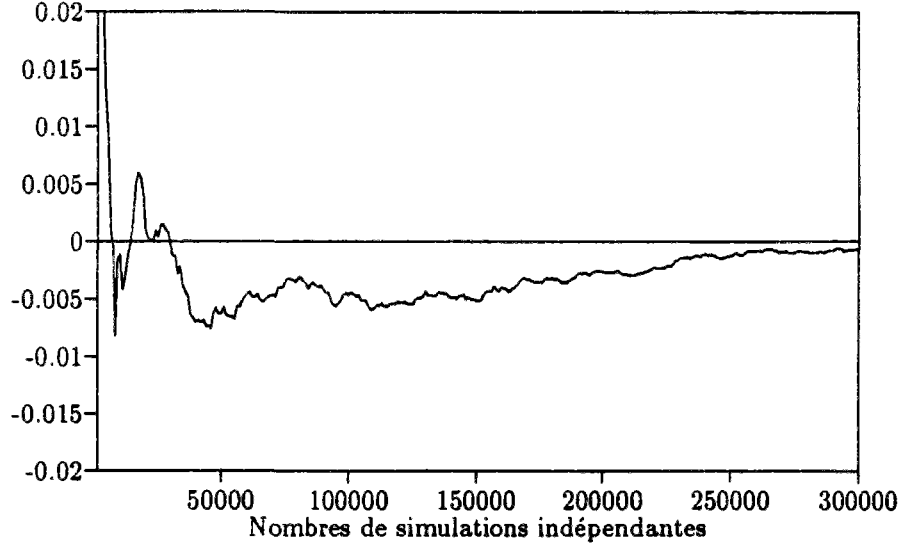


Figure 2.2: Calcul par Monte Carlo de  $E(|B_{\tau_1}|^2 - d\tau_a)$

On vérifie alors expérimentalement cette propriété par Monte-Carlo. Pour cela on fait apparaître dans la figure 2.2 l'espérance approchée de  $E(|B_{\tau_a}|^2 - d\tau_a)$ . Ces résultats ont été obtenus en dimension 3 avec  $a = 1$ .

On constate expérimentalement que le nombre aléatoire d'appels à la fonction random pour simuler le couple  $(\tau_a, B_{\tau_a})$  est de l'ordre de  $8d$ .

Le temps de calcul d'une simulation du couple reste négligeable pour nos applications, de l'ordre de  $10^{-3}d$  cpu sur un Spark 2 soit environ (annexe1)  $35d$  fois le temps de génération d'un nombre pseudo-aléatoire.

Voir les remarques en 2.2 concernant une éventuelle accélération de la simulation.

## 2.4 Etude du temps de sortie $\tau_{a,\mu}$ d'un cube pour un brownien de drift $\mu$ .

Soit  $(B_t)$  un brownien réel issu de zéro. On note pour  $a > 0$ , et  $\mu \in \mathbb{R}$ ,

$$\tau_{a,\mu} = \inf\{t > 0, |B_t + \mu t| = a\}.$$

On note aussi  $T_{a,\mu}$  le temps d'arrêt défini par

$$T_{a,\mu} = \inf\{t > 0, B_t + \mu t = a\},$$



avec la convention  $T_{a,0} = T_a$ .

Nous donnons quelques propriétés de  $\tau_{a,\mu}$  et  $T_{a,\mu}$  qui nous seront utiles pour étudier un schéma de discrétisation spatial pour les E.D.S., puis nous proposons une méthode de simulation numérique du couple  $(\tau_{a,\mu}, B_{\tau_{a,\mu}})$ .

#### 2.4.1 Quelques propriétés de $\tau_{a,\mu}$ et $T_{a,\mu}$ .

On va appliquer le théorème de Girsanov pour trouver des propriétés sur  $\tau_{a,\mu}$ . Pour cela on définit le processus  $B_t^\mu$  par

$$B_t^\mu = B_t - \mu t$$

et la martingale exponentielle  $Z_t$  par

$$Z_t = \exp(\mu B_t - \frac{1}{2}\mu^2 t).$$

On définit la probabilité  $\mathbf{P}^\mu$  par  $\frac{d\mathbf{P}^\mu}{d\mathbf{P}} = Z$ . Sous la probabilité  $\mathbf{P}^\mu$ ,  $B_t^\mu$  est un mouvement brownien,  $B_t$  un mouvement brownien de drift  $\mu$  et  $\tau_a$  le temps de sortie de  $B_t$ .

On note  $f_{\tau_{a,0}}$  la densité de  $\tau_a$  sous  $\mathbf{P}$  (donnée en 2.2.1 Proposition 2). On a la

**Proposition 8.** *Sous la probabilité  $\mathbf{P}$  on a les propriétés*

$$\forall x < \mu^2/2, \mathbf{E}(\exp(x\tau_{a,\mu})) = \frac{ch(\mu a)}{ch(a\sqrt{\mu^2 - 2x})}, \quad (2.3)$$

$$\mathbf{E}(\tau_{a,\mu}) = \frac{a}{\mu} th(\mu a), \quad (2.4)$$

$$\forall p \geq 0, \mathbf{E}(\tau_{a,\mu}^p) \leq C_p \min\{a^{2p}, (\frac{a}{|\mu|})^p\} \quad (2.5)$$

$$\int_0^\infty \mathbf{P}(\tau_{a,\mu} > s)^{1/2} s^p ds \leq C_p \min\{(\frac{a}{|\mu|})^{p+1}, a^{2p+2}\}, \quad (2.6)$$

$$\mathbf{P}(T_{a,\mu} < T_{-a,\mu}) = \frac{1}{1 + \exp(-2\mu a)}, \quad (2.7)$$

De plus, si  $f_{\tau_{a,\mu}}(t)$  désigne la densité de  $\tau_{a,\mu}$  sous  $\mathbf{P}$ , alors on a

$$f_{\tau_{a,\mu}}(t) = ch(\mu a) \exp(-\frac{1}{2}\mu^2 t) f_{\tau_{a,0}}(t). \quad (2.8)$$

$f_{\tau_{a,0}}(t)$  étant donnée en 2.2.1 proposition 2.

Ces propriétés sont sans doute classiques mais nous ne les avons pas trouvées dans la littérature, sauf (2.3) pour  $x < 0$ , et 2.7, qui apparaissent par exemple dans Kent[78]. Nous donnons donc des preuves complètes.

**Preuve:** Montrons (2.3). Fixons un réel positif  $t$ . La v.a  $\exp(x(\tau_a \wedge t))$  étant intégrable, par Girsanov, il vient

$$\mathbf{E}^\mu(\exp(x(\tau_a \wedge t))) = \mathbf{E}(Z_t \exp(x(\tau_a \wedge t))) = \mathbf{E}(\exp(x(\tau_a \wedge t)) \mathbf{E}(Z_t | \mathcal{F}_{\tau_a \wedge t}))$$

$$\mathbf{E}^\mu(\exp(x(\tau_a \wedge t))) = \mathbf{E}(Z_{(\tau_a \wedge t)} \exp(x(\tau_a \wedge t))) = \mathbf{E}(\exp(\mu B_{(\tau_a \wedge t)} - \frac{1}{2} \mu^2 (\tau_a \wedge t)) \exp(x(\tau_a \wedge t))).$$

En remarquant que  $\exp((x - 1/2 \mu^2) \tau_a)$  est intégrable sous  $\mathbf{P}$  par (iii) de 2.2.1 Proposition 1, le membre de droite converge par convergence dominée. Par convergence monotone le membre de gauche converge. On peut ainsi passer à la limite. Enfin  $B_{\tau_a}$  est indépendant de  $\tau_a$  et prend les valeurs  $a$  et  $-a$  de façon équiprobable, il vient donc

$$\mathbf{E}^\mu(\exp(x \tau_a)) = ch(\mu a) \mathbf{E}(\exp((x - \frac{1}{2} \mu^2) \tau_a)).$$

La proposition 1 (i) de 2.2.1 donne alors le résultat. Pour montrer (2.4) on remarque que

$$ch(a\sqrt{\mu^2 - 2x}) = ch(a\mu) - \frac{ax}{\mu} sh(\mu a) + o(x),$$

d'où

$$\mathbf{E}^\mu(\exp(x \tau_a)) = 1 + \frac{ax}{\mu} th(\mu a) + o(x),$$

qui prouve le résultat. Pour (2.5) on utilise Girsanov,

$$\mathbf{E}^\mu(\tau_a^p) = \mathbf{E}(\exp(\mu B_{\tau_a} - \frac{1}{2} \mu^2 \tau_a) \tau_a^p),$$

puis comme précédemment on obtient,

$$\mathbf{E}^\mu(\tau_a^p) = ch(\mu a) \mathbf{E}(\tau_a^p \exp(-\frac{1}{2} \mu^2 \tau_a)).$$

D'après la proposition 1 (i) de 2.2.1

$$\mathbf{E}^\mu(\tau_a^p) = ch(\mu a) a^{2p} \mathbf{E}(\tau_1^p \exp(-\frac{1}{2} \mu^2 a^2 \tau_1)). \quad (2.9)$$

On a alors d'après la proposition 1.(ii) de 2.1.1,

$$\mathbf{E}(\tau_1^p \exp(-\frac{1}{2} \mu^2 a^2 \tau_1) \mathbf{1}_{\{\tau_1 \leq 1\}}) \leq \mathbf{E}(\exp(-\frac{1}{2} \mu^2 a^2 \tau_1)) = \frac{1}{ch(\mu a)},$$

et

$$\mathbf{E}(\tau_1^p \exp(-\frac{1}{2} \mu^2 a^2 \tau_1) \mathbf{1}_{\{\tau_1 > 1\}}) \leq \exp(-\frac{1}{2} \mu^2 a^2) \mathbf{E}(\tau_1^p).$$

Donc d'après (2.9),

$$\mathbf{E}^\mu(\tau_a^p) \leq a^{2p} (1 + C_p \exp(\mu a - \frac{1}{2} \mu^2 a^2)).$$

En remarquant que  $\exp(\mu a - \mu^2 a^2 / 2) \leq \exp(1/2)$ , il vient

$$\mathbf{E}^\mu(\tau_a^p) \leq C_p a^{2p}.$$

Distinguons maintenant deux cas.

- Si  $|\mu|a \leq 1$ , l'inégalité précédente permet d'écrire

$$\mathbf{E}^\mu(\tau_a^p) \leq C_p \left(\frac{a}{|\mu|}\right)^p.$$

- Si  $|\mu|a > 1$ , on a d'après 2.2.1 propriété 1.(ii):

$$\mathbf{E}(\tau_1^p \exp(-\frac{1}{2}\mu^2 a^2 \tau_1) \mathbf{1}_{\{\tau_1 \leq \frac{4}{|\mu|a}\}}) \leq \frac{4^p}{|\mu|^p a^p} \mathbf{E} \exp(-\frac{1}{2}\mu^2 a^2 \tau_1) = \frac{C_p}{|\mu|^p a^p \operatorname{ch}(\mu a)}.$$

Puis

$$\mathbf{E}(\tau_1^p \exp(-\frac{1}{2}\mu^2 a^2 \tau_1) \mathbf{1}_{\{\tau_1 > \frac{4}{|\mu|a}\}}) \leq \exp(-|\mu|a) \mathbf{E}(\tau_1^p \exp(-\frac{1}{4}\mu^2 a^2 \tau_1) \mathbf{1}_{\{\tau_1 > \frac{4}{|\mu|a}\}}).$$

En utilisant

$$\forall x > 0, x^p \exp(-\lambda x) \leq p^p e^{-p} \lambda^{-p},$$

on obtient avec  $x = \tau_1$ ,  $\lambda = \mu^2 a^2 / 4$ ,

$$\mathbf{E}(\tau_1^p \exp(-\frac{1}{2}\mu^2 a^2 \tau_1) \mathbf{1}_{\{\tau_1 > \frac{4}{|\mu|a}\}}) \leq C_p \exp(-|\mu|a) \frac{1}{\mu^{2p} a^{2p}} \leq C_p \exp(-|\mu|a) \frac{1}{|\mu|^p a^p}.$$

On a ainsi d'après (2.9)

$$\mathbf{E}^\mu(\tau_a^p) \leq \left(\frac{a}{|\mu|}\right)^p,$$

ce qui prouve (2.5). Pour montrer (2.6) on écrit que

$$\mathbf{P}^\mu(\tau_a > s) \leq \frac{\mathbf{E}^\mu(\tau_a^{2p+4})}{s^{2p+4}} \leq C \left(\frac{a}{|\mu|}\right)^{2p+4} \frac{1}{s^{2p+4}},$$

$$\int_{\frac{a}{|\mu|}}^{+\infty} \mathbf{P}^\mu(\tau_a > s)^{1/2} s^p ds \leq C \left(\frac{a}{|\mu|}\right)^{p+1}.$$

$$\int_0^{\frac{a}{|\mu|}} \mathbf{P}^\mu(\tau_a > s)^{1/2} s^p ds \leq \int_0^{\frac{a}{|\mu|}} s^p ds \leq C \left(\frac{a}{|\mu|}\right)^{p+1}.$$

De même on a

$$\mathbf{P}^\mu(\tau_a > s) \leq \frac{\mathbf{E}^\mu(\tau_a^{2p+4})}{s^{2p+4}} \leq C a^{2(2p+4)} \frac{1}{s^{2p+4}},$$

$$\int_{a^2}^{+\infty} \mathbf{P}^\mu(\tau_a > s)^{1/2} s^p ds \leq C a^{2p+2}.$$

$$\int_0^{a^2} \mathbf{P}^\mu(\tau_a > s)^{1/2} s^p ds \leq \int_0^{a^2} s^p ds \leq C a^{2p+2}.$$

Ce qui démontre la formule intégrale de (2.6). Pour montrer (2.7) on écrit que

$$\begin{aligned} \mathbf{E}^\mu(1_{\{T_a < T_{-a}\}}) &= \mathbf{E}(1_{\{T_a < T_{-a}\}} Z_{T_a}) \\ &= \exp(\mu a) \mathbf{E}(1_{\{T_a < T_{-a}\}} \exp(-\frac{1}{2}\mu^2 T_a)) \end{aligned}$$

et d'après Itô-McKean[65] 1.7 pb 6,

$$\mathbf{E}(1_{\{T_a < T_{-a}\}} \exp(-\frac{1}{2}\mu^2 T_a)) = \frac{1}{2 \operatorname{ch}(\mu a)}.$$

On en déduit alors (2.7). Enfin pour montrer (2.8) on remarque que

$$\mathbf{P}^\mu(\tau_a \leq t) = \mathbf{E}(1_{\{\tau_a \leq t\}} Z_{\tau_a}) = \mathbf{E}(1_{\{\tau_a \leq t\}} \exp(\mu B_{\tau_a} - \frac{1}{2}\mu^2 \tau_a)).$$

Et comme précédemment on obtient

$$\mathbf{P}^\mu(\tau_a \leq t) = \operatorname{ch}(\mu a) \mathbf{E}(1_{\{\tau_a \leq t\}} \exp(-\frac{1}{2}\mu^2 \tau_a)),$$

qui donne le résultat.

□

#### 2.4.2 Simulation de $(\tau_{a,\mu}, B_{\tau_{a,\mu}})$ .

D'après la propriété précédente, on connaît la densité de  $\tau_{a,\mu}$ . Cette densité est très semblable à celle de  $\tau_{a,0}$  étudiée en 2.2.2. La simulation se fait de manière analogue en incluant  $\operatorname{ch}(\mu a)$  dans la constante de domination et en remarquant que  $\exp(-1/2 \mu^2 t) \leq 1$ .

Remarquons que, dans l'utilisation que nous ferons de cet algorithme au chapitre 6, nous nous placerons toujours dans le cas où  $\mu a \leq 1$ . Si cette condition n'est pas vérifiée, alors le nombre de rejets peut devenir important, nuisant à la rapidité de la simulation. Il faudrait alors étudier une majoration plus fine de la densité si tel était le cas.

La formule (2.7) de la propriété précédente permet de simuler  $B_{\tau_{a,\mu}}$  à partir de  $\tau_{a,\mu}$  car

- Si  $T_a < T_{-a}$  alors  $B_{\tau_{a,\mu}} = a - \mu \tau_{a,\mu}$ .
- Si  $T_a > T_{-a}$  alors  $B_{\tau_{a,\mu}} = -a - \mu \tau_{a,\mu}$ .

On peut ainsi simuler le couple  $(\tau_{a,\mu}, B_{\tau_{a,\mu}})$ .

Sur la courbe de la figure 2.3 apparaît l'erreur entre la valeur approchée de  $\mathbf{E}\tau_{1,1}$  calculée sur  $N$  simulations indépendantes par Monte Carlo, et la valeur exacte (donnée par (2.4)) et qui est de l'ordre de l'unité.

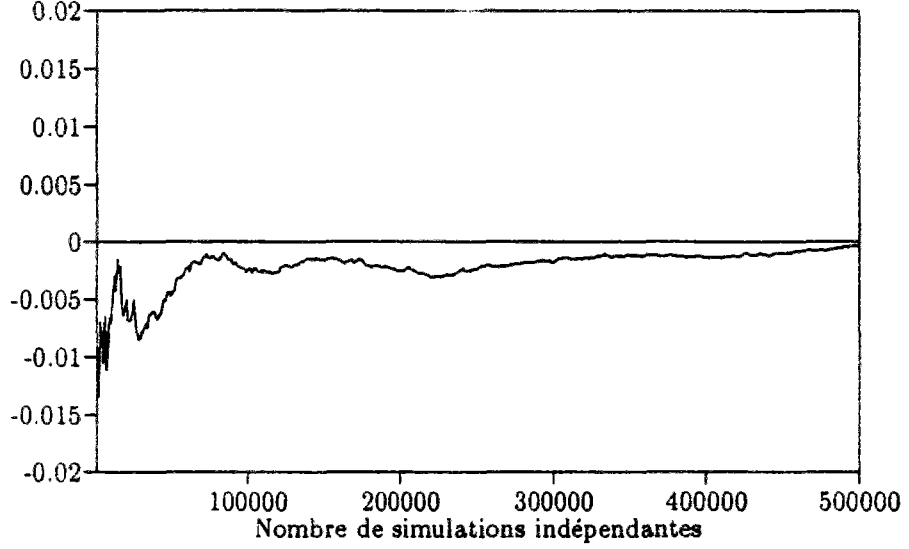


Figure 2.3: Calcul par Monte Carlo de l'erreur précédente

### 2.4.3 Cas de la dimension $d$ .

Une généralisation naturelle est de considérer

$$\tau = \inf\{t \geq 0, |\sigma W_t + \mu t| = a\},$$

où  $\sigma$  est une matrice  $\mathbf{R}^{n \times d}$ ,  $\mu$  un vecteur de  $\mathbf{R}^n$  et  $a$  un réel positif.  $W_t$  désigne un brownien  $d$  dimensionnel issu de zéro, et  $|\cdot|$  désigne la norme du max.

Un tel temps de sortie interviendra dans le chapitre sur la discrétisation spatiale des diffusions. Nous donnons une propriété de majoration de la norme  $L^p$  de  $\tau$  qui nous sera alors utile.

Nous noterons  $\sigma_o$  et  $\mu_o$  deux réels vérifiant

$$0 \leq \sigma_o \leq \sqrt{\sum_{i,j} \sigma_{i,j}^2} \text{ et } 0 \leq \mu_o \leq \sqrt{\sum_i \mu_i^2}.$$

et nous utiliserons la convention  $1/x = +\infty$  si  $x = 0$ . On a alors la

**Proposition 9.**

$$\forall p \geq 0, (\mathbf{E}\tau^p)^{1/p} \leq C_p \min\left\{\frac{a^2}{\sigma_o^2}, \frac{a}{\mu_o}\right\},$$

$$\int_0^\infty \mathbf{P}(\tau > s)^{1/2} s^p \leq C_p \min\left(\frac{a^2}{\sigma_o^2}, \frac{a}{\mu_o}\right)^{p+1}.$$

**Preuve:** On peut supposer que  $\sigma_o$  et  $\mu_o$  sont tous deux non nuls. Posons  $\lambda_i = \sqrt{\sum_j \sigma_{i,j}^2}$ . Il existe un indice  $i_o$  pour lequel on a  $\lambda_{i_o} \geq \sigma_o/\sqrt{d} > 0$ . Notons alors

$$B_t^i = \frac{1}{\lambda_i} \sum_j \sigma_{i,j} W_t^j,$$

et

$$\tau_i = \inf\{t \geq 0, |\sum_j \sigma_{i,j} W_t^j + \mu_i t| = a\}.$$

Il est clair que  $B_t^{i_o}$  est un brownien standard et que  $\tau \leq \tau_{i_o}$ . De plus

$$\tau_{i_o} = \inf\{t \geq 0, |\lambda_{i_o} B_t^{i_o} + \mu_{i_o} t| = a\},$$

donc d'après la proposition 8 de 2.4.1 on a

$$\|\tau\|_{L^p} \leq \|\tau_{i_o}\|_{L^p} \leq C_p \frac{a^2}{\sigma_o^2} d.$$

On procède de même pour obtenir le terme en  $\frac{a}{\mu_o}$  et la majoration de l'intégrale.

□

**Simulation de  $(\tau, W_\tau)$ .** Nous ne savons pas traiter le cas  $\mu \neq 0$  et même dans le cas où  $\mu = 0$ , nous ne savons pas simuler  $W_\tau$ , sauf dans le cas où la matrice  $\sigma$  est diagonale: L'indépendance des composantes était en effet nécessaire en 2.3 pour faire la simulation.

On pourrait alors penser simuler le *temps de sortie d'une sphère* et non d'un cube, car la distribution du brownien avec drift sur la sphère est connue. Mais cette fois nous ne savons pas simuler "exactement" (voir 2.5.3) le temps d'atteinte de la sphère et ceci même si  $\mu = 0$ .

## 2.5 Simulation du brownien à sa sortie d'un ouvert borné.

Nous présentons maintenant une application possible des algorithmes précédents sur un cas d'école: Faire une simulation du brownien à sa sortie d'un ouvert borné.

Soit  $U$  un ouvert borné de  $\mathbb{R}^d$ . On supposera, sans restreindre notre étude, que  $U$  contient l'origine et on notera  $B_t$  un brownien  $d$  dimensionnel issu de zéro.

On notera  $\partial U$  la frontière de  $U$ , et  $d(x,y)$  désignera la distance euclidienne de  $x$  à  $y$ . On note

$$\eta = \inf\{t > 0, B_t \notin U\}.$$

Notre objectif est de proposer une approximation p.s du couple  $(\eta, B_\eta)$ .

On peut remarquer qu'il existe des méthodes déterministes pour calculer des approximations de la loi de  $(\eta, B_\eta)$ . En effet cette loi est la mesure harmonique du point  $(0,0)$  pour le processus  $(t, B_t)$  (Bauer[66]). Mais la méthode trajectorielle présentée ici s'accorde à l'étude des E.D.S conduites par un mouvement brownien.

Donnons pour commencer une propriété très générale sur la régularité des temps de sortie.

Soit  $(S, d)$  un espace métrique et  $F$  un fermé de  $S$ . On définit pour  $f \in C(R^+, S)$  le temps d'entrée dans  $F$  par

$$T(f) = \inf\{t \geq 0, f(t) \in F\},$$

et

$$T^+(f) = \inf\{t \geq 0, f(t) \in \overset{\circ}{F}\},$$

le temps d'entrée dans l'intérieur de  $F$ . On a alors la

**Proposition 10.** *Si l'espace  $C(R^+, S)$  est muni de la norme de la convergence uniforme, alors la fonctionnelle  $T$  est s.c.i.. Si de plus  $T^+(f) = T(f)$  alors  $T$  est continue en  $f$ .*

**Preuve:** Montrons que  $T$  est s.c.i.. Pour cela considérons une suite  $f_n$  qui converge uniformément vers  $f$ . Pour  $t < T(f)$ ,  $d(f(t), F) = \epsilon > 0$ . Par convergence uniforme

$$\exists n_0, \forall n > n_0, d(f_n(t), F) > \frac{\epsilon}{2},$$

et donc  $T(f_n) > t$ . Ainsi  $\liminf T(f_n) \geq t$  pour tous  $t < T(f)$ , et donc  $\liminf T(f_n) \geq T(f)$ .  $T$  est donc s.c.i..

Pour montrer la deuxième partie de la proposition montrons que si  $f_n$  converge vers  $f$  alors  $\limsup T(f_n) \leq T^+(f)$ . Pour cela considérons  $t$  tel que  $f(t) \in \overset{\circ}{F}$ . Comme précédemment il existe  $n_0$  tel que,  $\forall n > n_0$ ,  $f_n(t) \in \overset{\circ}{F}$ , donc  $t \geq T(f_n)$ , puis  $t \geq \limsup T(f_n)$ . Ainsi  $\limsup T(f_n) \leq T^+(f)$ . Ce qui prouve la continuité de  $T$  en  $f$  si on suppose  $T^+(f) = T(f)$ .

□

**Remarque 3** Cette démonstration s'inspire d'un résultat plus général de Pagès[89] qui donne la continuité au sens de Skorokhod pour  $f \in D(R^+, S)$ . On peut aussi consulter Jacod-Shiryaev[87] ou Pagès[86] pour un résultat identique avec  $T(f) = \inf\{t \geq 0, |f(t)| \text{ ou } |f(t^-)| > M\}$ .

### 2.5.1 Une méthode élémentaire de simulation .

L'idée la plus naturelle consiste à prendre un processus de Markov qui converge p.s vers une trajectoire brownienne (en fait ici une discrétisation itérative classique du mouvement brownien) et de montrer que le temps de sortie de ce processus converge p.s vers le temps de sortie du brownien.

Nous allons voir que ce résultat est vrai sous certaines hypothèses, mais que cette méthode ne semble pas pertinente dans la pratique, notamment dans le cas du cercle étudié plus loin.

On note  $T$  un réel positif fixé, et  $\Delta t$  le pas de temps de l'approximation et on pose

$$\eta^T = \min\{\eta, T\}.$$

On note aussi pour tout entier  $k$ ,  $t_k = k \Delta t$  avec  $\Delta t = \frac{T}{N}$  et on définit le processus  $(X^N)$  par

$$\forall t \in [t_k, t_{k+1}[ , X_t^N = B_{t_k}.$$

On définit aussi le temps de sortie du processus de l'ouvert  $U$  et on pose

$$\eta_N^T = \min\{ \inf\{t, X_t^N \notin U\}, T\}.$$

On a alors la

**Proposition 11.** *Si  $U$  est un ouvert borné de frontière lipschitzienne, alors  $\eta_N^T$  converge p.s vers  $\eta^T$  quand  $N$  tend vers l'infini.*

**Preuve :**

- Il y a convergence uniforme p.s de  $X_t^N$  vers le brownien sur le compact  $[0, T]$ . Ce résultat classique peut par exemple se démontrer en remarquant que:

$$\forall t \in [t_k, t_{k+1}], |X_t^N - B_t| \leq \sup_{t \in [t_k, t_{k+1}]} |B_t - B_{t_k}|$$

et ainsi pour tous  $p$ ,

$$\mathbf{E}(\sup_{t \in [0, T]} |X_t^N - B_t|^{2p}) \leq \mathbf{E}(\sup_{k=1..N} \sup_{t \in [0, \Delta t]} |W_t^k|^{2p})$$

où les  $W_t^k$  sont  $N$  browniens indépendants. Donc

$$\mathbf{E}(\sup_{t \in [0, T]} |X_t^N - B_t|^{2p}) \leq \sum_{k=1}^N \mathbf{E}(\sup_{t \in [0, \Delta t]} |W_t^k|^{2p}) \leq C \frac{T^p}{N^{p-1}}.$$

Il reste à choisir  $p > 2$ , la série de terme général  $\sup_{t \in [0, T]} |X_t^N - B_t|^{2p}$  converge alors dans  $L^1$ , et il y a donc p.s convergence uniforme sur  $[0, T]$  du terme général vers 0.

- Montrons qu'il y a convergence p.s de  $\eta_N^T$  vers  $\eta^T$ . La frontière de  $U$  étant lipschitzienne, elle vérifie la condition de cône de Zaremba et d'après Karatzas-Shreve[88] 4.2.C p.250, tous les points de la frontière sont alors réguliers (i.e  $T^+ = T$ ). En appliquant la proposition 10, la fonctionnelle  $T$  est continue pour la norme de la convergence uniforme en  $B_t$ , ce qui prouve le résultat.

□

**Remarque 4** *Cette méthode donne une approximation supérieure du temps de sortie de l'ouvert car certaines trajectoires browniennes sont déjà sorties entre deux temps de discrétisation  $t_k$  et  $t_{k+1}$ .*

Il faut noter que cette première approximation du temps de sortie présente deux défauts:

- On ne dispose pas d'estimation de la vitesse de convergence ni de critère d'arrêt, et on ne sait donc pas en pratique quelle erreur est commise lors de l'approximation, ni comment choisir le pas de discrétisation pour contrôler cette erreur.



- Lors d'une simulation, si l'on veut "améliorer" la précision de l'approximation, il faut retirer une nouvelle trajectoire du processus approximant.

Nous allons voir une autre méthode qui permet une approximation croissante du temps de sortie et l'utilisation d'un critère d'arrêt (spatial) puis nous comparerons ces deux méthodes en 2.5.3.

### 2.5.2 La méthode des "boîtes de contrôle".

L'idée est de construire la plus grande boîte noire, ici un cube centré en 0 et parallèle aux axes, qui soit contenue dans l'ouvert  $U$  et de simuler le brownien à la sortie de cette boîte, puis de réitérer cette procédure à partir du point de sortie précédent...

On introduit pour cela la chaîne de Markov suivante:

**Définition 2.** La suite  $(\eta_k, B_{\eta_k})$  est définie par

$$\eta_0 = 0.$$

$$\eta_{k+1} = \eta_k + \inf\{t > 0, \max_i |B_{\eta_k+t}^i - B_{\eta_k}^i| = a_k\}$$

avec

$$a_k = \frac{d(B_{\eta_k}, \partial U)}{(1 + \epsilon)\sqrt{d}}, \quad \epsilon > 0.$$

La constante  $\epsilon$  joue un rôle technique, elle permet d'éviter des problèmes de régularité à la frontière. En pratique elle sera prise très proche de 0 (et même identiquement nulle si l'ouvert est suffisamment régulier).

On a alors la propriété suivante qui assure la convergence de la procédure d'approximation:

**Proposition 12.** Soit  $U$  un ouvert borné de  $\mathbf{R}^d$  contenant l'origine. Si

$$\eta = \inf\{t > 0, B_t \notin U\}$$

désigne le temps de sortie de  $U$ , alors la suite croissante  $\eta_k$  converge p.s et dans  $L^p$  vers  $\eta$ , tout en restant inférieure.

**Preuve de la proposition 12:** L'ouvert étant borné, il vérifie les propriétés suivantes:

- $\forall p > 0, \mathbf{E}\eta^p < +\infty$  (et  $\eta < +\infty$  p.s ).
- $\forall x \in U$ , si  $d = d(x, \partial U)$  alors  $d < +\infty$  et  $B_o(x, d) \subset U$  où  $B_o(x, d)$  désigne la boule ouverte de centre  $x$  et de rayon  $d$  pour la norme euclidienne.

Ainsi, par récurrence sur  $k$ , si  $B_{\eta_k} \in U$ , alors

$$\forall t \in [\eta_k, \eta_{k+1}], d(B_{\eta_k}, B_t) \leq \sqrt{d} \max_i |B_{\eta_{k+1}}^i - B_{\eta_k}^i| < d(B_{\eta_k}, \partial U),$$

et donc

$$\forall k \geq 0, \forall t \in [0, \eta_k], B_t \in U.$$

En particulier on a

$$\forall k \geq 0, \eta_k < \eta. \quad (2.10)$$

$\eta$  étant p.s fini, la suite croissante  $(\eta_k)_k$  converge p.s et on a  $\limps(\eta_k) \leq \eta$ .

Montrons l'inégalité inverse. On remarque que:

$$\forall k \geq 0, \frac{d(B_{\eta_k}, B_{\eta_{k+1}})}{\sqrt{d}} \leq a_k \leq d(B_{\eta_k}, B_{\eta_{k+1}}). \quad (2.11)$$

La suite  $(\eta_k)$  étant p.s convergente et la trajectoire brownienne étant continue, on a

$$d(B_{\eta_k}, B_{\eta_{k+1}}) \xrightarrow{p.s} 0.$$

D'après (2.11) il y a donc convergence p.s de  $a_k$  vers zéro et on a donc

$$d(B_{\eta_k}, \partial U) \xrightarrow{p.s} d(B_{\limps(\eta_k)}, \partial U) = 0.$$

Ainsi  $B_{\limps(\eta_k)} \in \partial U$  et donc  $\limps(\eta_k) \geq \eta$ . Enfin il y a convergence dans  $L^p$  car  $\eta_k \leq \eta \in L^p$ , ce qui achève la démonstration.

□

**Remarque 5** La propriété précédente s'étend au cas d'un processus continu quelconque dont le temps de sortie de l'ouvert est p.s fini.

En théorie la suite  $(\eta_k, B_{\eta_k})$  converge donc vers le couple  $(\eta, B_\eta)$ , mais en pratique la suite  $(\eta_k)_k$  est a priori infinie. Il faut donc, pour la simulation numérique, utiliser un critère d'arrêt.

On ne dispose pas à priori de critère d'arrêt temporel, ni d'estimation de la vitesse de convergence, mais on conviendra que le temps de sortie est atteint quand la distance de  $B_{\eta_k}$  au bord est négligeable, i.e inférieure à une constante fixée  $\alpha$  très proche de zéro. On pose alors

$$\eta_\infty^\alpha = \inf\{\eta_k, d(B_{\eta_k}, \partial U) \leq \alpha\}.$$

Le couple  $(\eta_\infty^\alpha, B_{\eta_\infty^\alpha})$  est donc pris comme approximation de  $(\eta, B_\eta)$ . Ce critère d'arrêt est donc un critère d'arrêt spatial, tout se passe comme si on autorisait une déformation aléatoire infinitésimale de l'ouvert  $U$ , d'amplitude  $\alpha$ ,  $\omega$  par  $\omega$ .

Remarquons aussi que pour raffiner une approximation du temps de sortie il suffit de diminuer  $\alpha$ , et donc simplement de calculer plus de termes de la suite.

Dans le cas de la sphère on a une estimation précise de l'erreur commise. On a en effet la

**Proposition 13.** Soit  $S(a)$  la sphère (ouverte) de  $\mathbf{R}^d$  de rayon  $a$ . Si

$$\eta(a) = \inf\{t > 0, B_t \notin S(a)\}$$

désigne le temps de sortie de la sphère, alors la suite croissante  $\eta_k$  converge p.s et dans  $L^p$  vers  $\eta(a)$ , tout en restant inférieure. De plus la suite arrêtée  $(\eta_\infty^\alpha, B_{\eta_\infty^\alpha})$  vérifie les propriétés

$$\eta(a - \alpha) \leq \eta_\infty^\alpha \leq \eta(a),$$

et

$$0 \leq \mathbf{E}(\eta(a) - \eta_\infty^\alpha) \leq \frac{2\alpha a}{d}.$$

**Preuve de la proposition 13.** Il est clair que  $\eta(a - \alpha) \leq \eta_\infty^\alpha \leq \eta(a)$ . Il suffit ensuite de remarquer que  $\mathbf{E}(B_{\eta(a)}^2) = d\mathbf{E}(\eta(a)) = a^2$ , on a alors  $(a - \alpha)^2 \leq d\mathbf{E}(\eta_\infty^\alpha) \leq a^2$ , ce qui prouve la proposition.

□

Enfin pour la simulation numérique, il suffit d'utiliser les algorithmes de simulation des temps de sortie d'un cube pour le brownien qui ont été vus en 2.3. Il faut aussi pouvoir estimer la distance de tous points de l'ouvert à la frontière.

**Pourquoi ne pas simuler directement le temps de sortie d'une sphère ?** Il aurait été en effet plus simple d'étudier le temps de sortie d'une sphère car on sait que le brownien est alors uniformément distribué sus la sphère. Mais si

$$\tau_a = \inf\{t > 0, \|B_t\| = a\},$$

représente le temps d'atteinte de  $a$  pour un Bessel, nous ne connaissons pas de formule générale pour la densité de  $\tau_a$  qui permettrait de faire une simulation numérique. De Wendel[80] on a

$$\mathbf{E}(e^{-x\tau_a}) = \left(\frac{ax}{2}\right)^\nu \frac{1}{\Gamma(\nu)I_\nu(ax)}$$

où  $\nu = (d - 2)/2$  et  $I_\nu$  désigne la fonction de Bessel modifiée. Mais il reste ensuite à inverser cette transformé de Laplace. Précisons qu'une expression de la densité peut se trouver, pour la dimension  $d = 2$  dans Yor[80], pour la dimension  $d = 3$  dans Imhof[84] et pour les dimensions supérieures dans Pei Shu[85]. Le problème délicat, que nous n'avons pas résolu, est de trouver une écriture de ces densités permettant de les simuler.

**Remarque 6 .** Une méthode différente permet de simuler, de manière récursive, un brownien jusqu'à ce qu'il atteigne la sphère. On renvoie au cours de D.E.A de Bouleau[90a].

### 2.5.3 Application numérique au temps de sortie d'un cercle.

Plaçons-nous dans  $\mathbf{R}^2$ . Soit  $C$  le disque de centre 0 et de rayon 1. On désigne par

$$\tau = \inf\{t > 0, |B_t| \notin C\}$$

le temps de sortie du disque  $C$ . Nous allons utiliser les deux méthodes précédentes pour simuler une "bonne" approximation de  $(\tau, B_\tau)$  et vérifier que  $\mathbf{E}(\tau) = 1/2$ .

**Première méthode.** Le temps de sortie est pris comme le premier temps de sortie  $\tau^{\Delta t}$  de la chaîne de Markov approximante. Un minorant de ce temps de sortie est  $\tau^{\Delta t} - \Delta t$ .

Nous avons simulé, pour chaque valeur du paramètre  $\Delta t$ , 1 000 000 de trajectoires de la chaîne de Markov. Nous avons ainsi mesuré les statistiques suivantes:

- *Temps*, le temps moyen de calcul d'une trajectoire de  $\tau^{\Delta t}$  en temps cpu.
- *Dimension*, le nombre moyen d'appels à la fonction random pour simuler une trajectoire de  $\tau^{\Delta t}$ .
- *Erreur relative*, la quantité  $\mathbf{E} \frac{|\tau^{\Delta t} - 1/2|}{1/2}$ .

On est ici dans une situation typique où la dimension de la simulation, qui est ici d'environ 9000 pour le paramètre  $\Delta t = 0.0001$ , rend le calcul d'une espérance particulièrement délicat. Il aurait en effet fallu, pour calculer l'espérance précédente par Monte Carlo avec  $\Delta t = 0.0001$  et sur 1 000 000 de simulations, environ 2 semaines de calcul sur un Sun 4.

Par contre l'utilisation de la méthode du shift (voir Bouleau[90b] pour un exposé sur l'intégration en dimension grande et infinie et Ben Alaya[91] pour une étude détaillée du shift avec en particulier une estimation de la vitesse de convergence) permet de contourner cette difficulté.

Rappelons le principe de la méthode: Notons  $\theta$  le shift de Bernoulli défini sur  $[0, 1]^N$  par

$$\theta(U_1, \dots, U_n, \dots) = (U_2, \dots, U_{n+1}, \dots).$$

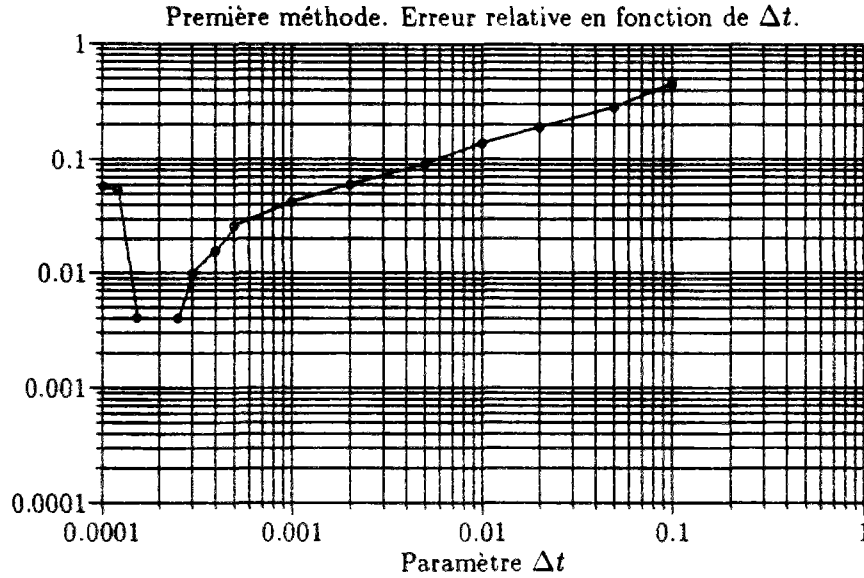
Si  $[0, 1]^N$  est muni de la mesure de Lebesgue produit  $\lambda$ , alors la transformation  $\theta$  est ergodique. Ainsi par le théorème ergodique, si  $F$  est  $\lambda$ -intégrable on a

$$\mathbf{E}F = \lim_N \frac{1}{N} \sum_{k=0}^{N-1} F \circ \theta^k.$$

Bien que l'on ne connaisse pas de résultat général sur la vitesse de convergence dans le théorème ergodique, on peut sous certaines hypothèses, ici  $\mathbf{E}(\tau^{\Delta t})^{2+\epsilon} < +\infty$ , obtenir (Ben Alaya[91]):

$$\forall \epsilon > 0, \left| \frac{1}{N} \sum_{k=0}^{N-1} F \circ \theta^k - \mathbf{E}F \right| = o\left(\frac{1}{\sqrt{N}} (\log N)^{3/2+\epsilon}\right).$$

On obtient ainsi une convergence très similaire à celle de la loi des grands nombres. L'intérêt est de diminuer grandement le temps de calcul de l'espérance et de réduire considérablement le nombre d'appels au générateur pseudo-aléatoire.



La première courbe représente l'erreur relative en fonction du paramètre  $\Delta t$ . Les valeurs ont été obtenues par la méthode du shift (et aussi par Monte Carlo pour  $\Delta t \leq 0.001$ ).

**Commentaires:** On constate que, jusqu'à  $\Delta t \leq 0.0005$ , on obtient une approximation supérieure du temps de sortie. Il est en effet prévisible que la trajectoire brownienne soit déjà sortie entre deux points de discrétisation. De plus la courbe précédente montre que pour ces valeurs de  $t$ :

$$E(\tau^{\Delta t} - \frac{1}{2}) \simeq O(\sqrt{\Delta t}).$$

C'est un phénomène plutôt surprenant car on aurait pu s'attendre à une erreur en  $\Delta t$  et non en  $\sqrt{\Delta t}$  provenant du dernier terme de la suite. Il semble donc que l'on ne puisse pas négliger le fait qu'une trajectoire soit sortie puis revenue dans l'ouvert entre deux points de discrétisation.

Pour  $\Delta t \leq 0.0005$ , nous avons obtenu des valeurs *inférieures* du temps de sortie. Enfin la précision diminue pour  $\Delta t \leq 0.0002$ . Ces résultats sont sans doute dus à des phénomènes d'instabilité numérique.

De plus pour des valeurs plus petites de  $\Delta t$ , la dimension d'une simulation et aussi par conséquent le temps de calcul deviennent prohibitifs, rendant illusoire l'utilisation d'une telle méthode.

Les statistiques concernant la dimension d'une simulation et les temps de calculs sont présentées plus loin pour permettre une comparaison avec la deuxième méthode.

**Utilisation de la deuxième méthode.** Nous utilisons la méthode des cubes de contrôle décrite en 2.5.2 pour approcher p.s le couple  $(\tau, B_\tau)$ .

On simule donc la chaîne de Markov  $(\eta_k, B_{\eta_k})$  définie par  $\eta_0 = 0$ ,  $B_{\eta_0} = 0$

$$\eta_{k+1} = \eta_k + \inf\{t > 0, \max_{i=1,2} |B_{\eta_k+t}^i - B_{\eta_k}^i| = a_k\}$$

avec

$$a_k = \frac{d(B_{\eta_k}, Fr(C))}{\sqrt{2}}.$$

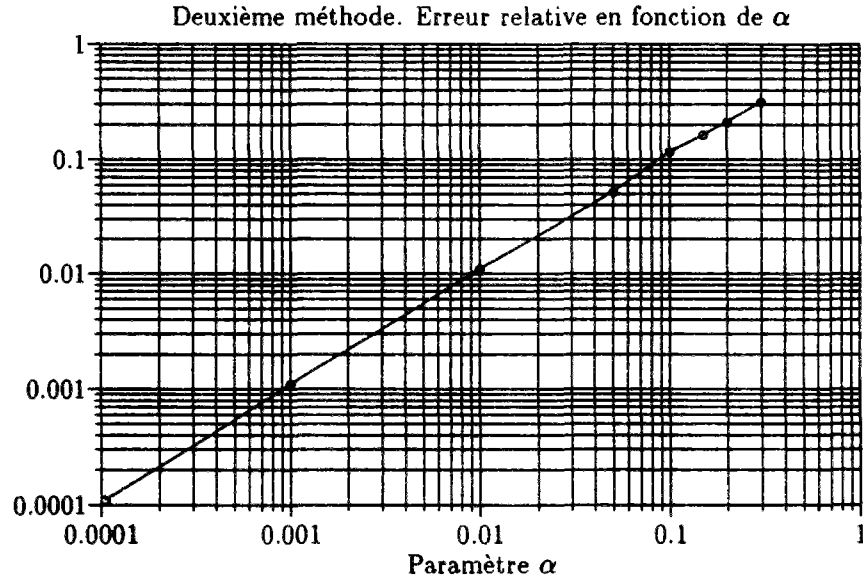
On considère que l'on a atteint la frontière du disque quand

$$d(B_{\eta_k}, Fr(C)) \leq \alpha.$$

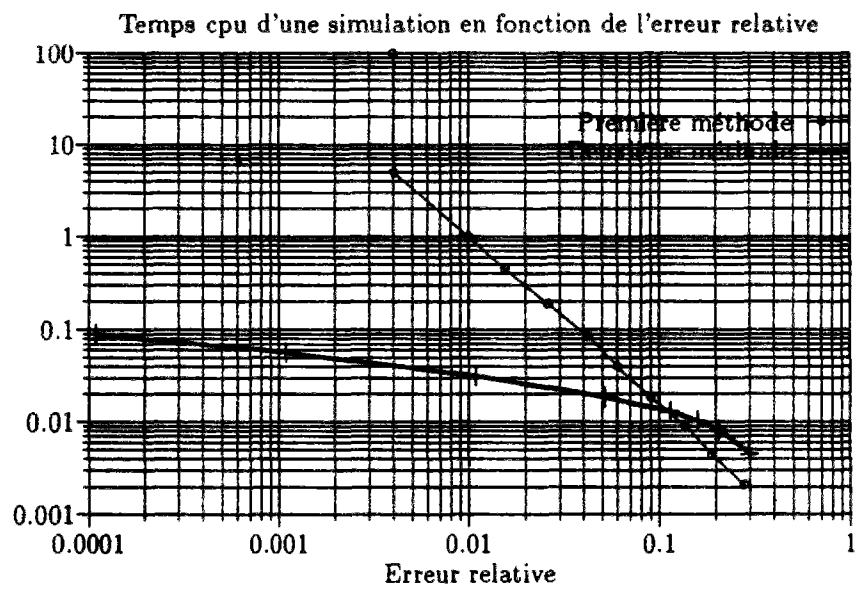
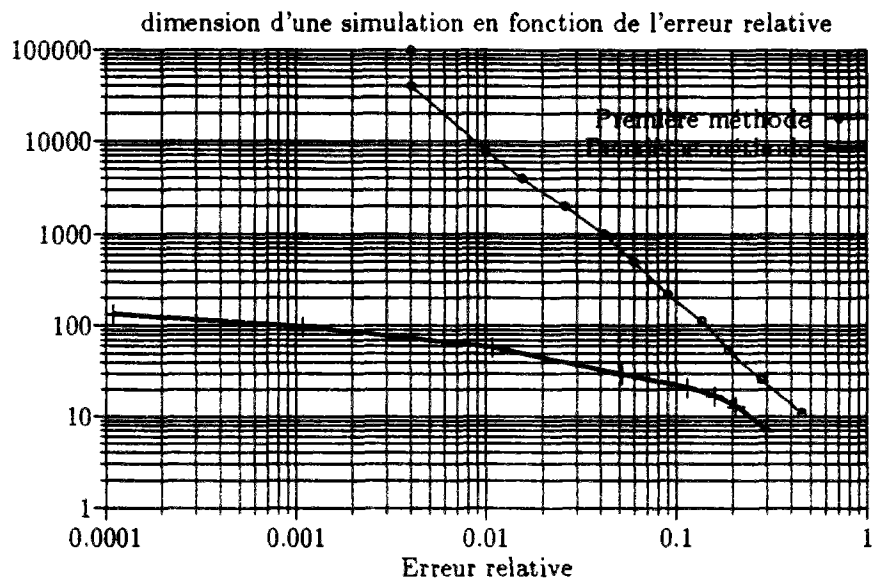
Ce critère d'arrêt donne une valeur approchée de  $\tau$  notée  $\tau_\infty$  qui vérifie d'après la proposition 13:

$$E(\tau_\infty - \frac{1}{2}) \in [-\alpha, 0].$$

Comme précédemment, nous présentons des statistiques obtenues sur 1 000 000 de simulations. Remarquons qu'il suffit de diminuer  $\alpha$ , et donc de calculer plus de termes de la suite  $(\nu_k, B_{\nu_k})$ , pour améliorer la précision.



**Comparaisons des deux méthodes d'approximation.** La première courbe présente, pour chacune des méthodes, le temps d'une simulation du temps de sortie en fonction de la précision de l'approximation. La deuxième représente la dimension d'une simulation (définie plus haut) en fonction de la précision.



Il ressort de ces statistiques que:

- Pour une précision de l'ordre de 10% les deux méthodes sont comparables en performances.
- Pour une plus grande précision, par exemple 1%, la deuxième méthode est plus rapide et beaucoup moins coûteuse en nombres pseudo aléatoires.
- Pour une précision plus grande, *seule la deuxième méthode est utilisable*.

Il semble donc que seule la méthode des "boîtes de contrôle" permette une approximation très fine du temps de sortie. Même si seule la loi d'un temps de sortie est recherchée, il semble qu'une approximation *trajectorielle* comme ici permette d'obtenir de bons résultats sans pour autant nuire à la rapidité de calcul.

## 2.6 Conclusion.

Nous avons proposé une méthode de simulation exacte du brownien à sa sortie d'un cube en 2.3 et testé sa validité numériquement. Puis nous avons montré sur un cas d'école, l'approximation du temps de sortie d'un cercle pour un brownien, comment utiliser cette méthode en contrôlant spatialement l'erreur d'approximation.

On peut aussi penser appliquer 2.5 à des problèmes d'arrêt optimal, ou d'enveloppe de Snell pour le brownien. En effet le temps d'arrêt optimal peut se représenter comme le temps de sortie d'un ouvert dans certains cas (voir Oksendal[85] par exemple).

Plus généralement nous pensons que cette méthode de simulation spatialement contrôlée du brownien permet de mieux rendre compte de sa *structure fine*. Elle semble à la fois plus naturelle et plus pertinente pour certaines applications que la traditionnelle méthode de simulation itérative exposée au chapitre 1. Nous verrons en particulier une utilisation de cette méthode pour simuler de façon fine des diffusions en 6.2 et 6.3.

On peut imaginer des extensions analogues comme par exemple:

- Faire un pont brownien conditionnellement à la non sortie d'un cube. Cela permettrait de faire ensuite des simulations récursives en certains points pour raffiner une trajectoire déjà tirée.
- Traiter complètement le cas de la dimension  $d$  en 2.4.3 qui permettrait d'étendre des résultats sur la simulation des E.D.S que nous verrons en 6.3. Etudier aussi le cas d'un cube décentré afin de simuler une diffusion sur un maillage donné par avance.
- Simuler la saucisse de Wiener pour contrôler les trajectoires du brownien jusqu'à l'infini.
- Simuler l'angle de rotation du brownien en dimension 2 pour observer des propriétés d'enroulement.





## Chapitre 3

# Simulation des diffusions: Introduction.



## 3.1 Qu'est ce qu'une équation différentielle stochastique?

### 3.1.1 Un exemple introductif.

Afin de motiver les chapitres suivants, revenons aux travaux de Langevin[1908] et considérons une particule sphérique de rayon  $a$  et de masse  $m$  en suspension dans un liquide de viscosité  $\nu$ . Si  $v_t$  désigne la vitesse de la particule, elle est soumise à une force de frottement  $-\nu v_t$  ainsi qu'à un très grand nombre de chocs moléculaires de résultante notée  $\sigma \xi_t$ . On a ainsi

$$\dot{v}_t = -\nu v_t + \sigma \xi_t. \quad (3.1)$$

Pour décrire complètement le mouvement de la particule il faut d'avantage de renseignements sur  $\xi_t$ . Les collisions étant symétriques, très nombreuses et presque instantanées, on peut supposer que  $\xi_s$  et  $\xi_t$  sont en moyenne nuls et non corrélés. Réécrivons (3.1) en introduisant la quantité  $B_t$  définie heuristiquement par  $\xi_t dt = B_{t+dt} - B_t$ . On a alors

$$dv_t = -\nu v_t + \sigma(B_{t+dt} - B_t). \quad (3.2)$$

Un modèle probabiliste naturel est donc de supposer que  $B_{t+dt} - B_t$  est une variable aléatoire centrée et indépendante de  $B_{s+ds} - B_s$  pour  $s < s + ds < t$ . Supposons de plus que  $B_t$  est continue en  $t$  puisque les chocs sont très nombreux. Il se trouve que le seul processus stochastique vérifiant ces propriétés est le mouvement brownien comme nous l'avons vu au chapitre 1. On peut alors par des considérations physiques obtenir les valeurs de  $\nu$  et  $\sigma$ . Ce modèle du mouvement d'une particule est connu sous le nom de modèle de Langevin.

Plus généralement considérons un système physique régit par une équation différentielle mais dont la dynamique est perturbée par un bruit aléatoire et qui peut s'écrire sous la forme

$$dX_t = f(t, X_t)dt + \sigma(t, X_t)\xi_t. \quad (3.3)$$

Il est toujours possible de faire l'hypothèse que le bruit en moyenne n'affecte pas le comportement de notre système, au besoin en modifiant la fonction  $f$ . Si de plus on peut supposer que les fluctuations du bruit sont rapides et non corrélées, ou encore que  $\xi_t$  est un "bruit neutre", alors on dit que  $\xi_t$  est un *bruit blanc*.

Un bruit blanc est défini dans la littérature appliquée comme un processus gaussien continu de densité spectrale constante (comme la lumière blanche), mais un tel processus n'existe pas mathématiquement! On aimerait écrire  $\xi_t = \dot{B}_t$  d'après ce qui précède, mais on a vu au chapitre 1 que le brownien est loin d'être dérivable. On peut néanmoins donner un sens à cette écriture, le bruit blanc est défini comme la dérivée au sens des distributions du brownien (voir Arnold[74]). Nous retiendrons simplement que *formellement* on remplacera toujours le bruit blanc  $\xi_t$  par le brownien  $B_t$  grâce aux relations heuristiques:

$$\xi_t dt = dB_t \text{ ou encore } B_t = \int_0^t \xi_s ds. \quad (3.4)$$

On est ainsi amené à écrire une généralisation de (3.2) et (3.3) à savoir:

$$dX_t = f(t, X_t)dt + \sigma(t, X_t)dB_t \quad (3.5)$$

comme modèle de comportement de notre système physique perturbé par un bruit blanc.

Nous allons maintenant donner un sens précis à (3.5), puis nous reviendrons ensuite en détail en 3.1.3 à notre problème de modélisation.

### 3.1.2 Propriétés classiques des diffusions.

Une équation différentielle stochastique (E.D.S) de type Itô est notée:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad t \in [0, T], \quad (3.6)$$

avec  $b: \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ ,  $\sigma: \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^{n \times d}$  et  $(W_t)$  un brownien standard de  $\mathbf{R}^d$ .

C'est un objet mathématique bien connu et, depuis les travaux de Itô (Itô[51]), les E.D.S ont donné lieu à une très vaste littérature. On renvoie à Gard[88] et Arnold[74] pour une introduction et à Karatzas-Shreve[88], Ikeda-Watanabe[81], pour une étude beaucoup plus complète. Essayons en quelques mots de rappeler comment interpréter l'écriture (3.6), puis dans quels cas et en quels sens il peut y avoir existence et unicité d'une solution.

L'équation (3.6) s'écrit aussi

$$X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s. \quad (3.7)$$

Nous avons rappelé au chapitre 1 que  $W_t$  n'est pas à variation bornée, aussi la deuxième intégrale dans (3.7) n'est pas une intégrale au sens ordinaire mais une intégrale au sens de Itô:

Considérons une subdivision quelconque de  $[0, T]$  notée  $\{t_k^n\}_k$  de pas maximum  $\delta_n$  et une fonction  $f$  qui est  $\mathcal{F}_t$  adaptée et appartient à  $L^2$ . Alors l'intégrale stochastique de type Itô est définie par

$$\int_0^T f(\omega, s)dW_s = \lim_{\delta_n \rightarrow 0} \sum_k f(\omega, t_k^n)(W_{t_{k+1}^n} - W_{t_k^n}), \quad (3.8)$$

la limite étant au sens  $L^2$ . L'intégrale au sens de Itô vérifie alors

$$\mathbf{E} \left( \int_0^T f(\omega, s)dW_s \right)^2 = \mathbf{E} \int_0^T f^2(\omega, s)ds.$$

De plus elle ne suit pas les règles du calcul différentiel ordinaire, à cause de la variation quadratique non nulle du brownien, et on a pour une fonction  $f$  de classe  $C^2$  la célèbre formule d'Itô, donnée dans le cas  $n = d = 1$  pour simplifier:

$$f(X_t) = f(X_s) + \int_s^t b(u, X_u)\partial_x f(X_u)du + \int_s^t \sigma(u, X_u)\partial_x f(X_u)dW_u + \frac{1}{2} \int_s^t \sigma^2(u, X_u)\partial_{xx}^2 f(X_u)du.$$

Dans quels cas l'équation (3.6) admet-elle une solution unique? La situation est plus complexe que pour les équations différentielles ordinaires. En effet on peut donner deux sens à la notion d'existence d'une solution:

- Il y a *existence et unicité d'une solution forte (ou trajectorielle)* si heuristiquement à une trajectoire du brownien  $W_t$  correspond une seule trajectoire solution de (3.6) et définie sur  $[0, T]$ .

- Il y a *existence et unicité d'une solution faible (ou en loi)* si il existe un ou plusieurs browniens dont les solutions correspondantes  $X_t$ , définies sur des espaces de probabilité à priori différents, ont toutes la même loi.

On montre que le premier point entraîne le deuxième, et on connaît de nombreux résultats permettant de savoir si l'équation (3.6) admet des solutions et en quels sens. Pour la suite nous retiendrons simplement la propriété suivante:

*Soit  $X_0$  de loi donnée. Si les coefficients  $b$  et  $\sigma$  sont globalement lipschitziens et à croissance linéaire en  $x$ , le tout uniformément en  $t$ , alors il y a existence et unicité d'une solution forte pour (3.6).*

Nous donnerons en 4.1 une formalisation précise de ces hypothèses de régularité sur les coefficients de l'équation. Remarquons que cette propriété ne suffit pas la plupart du temps à couvrir les cas usuels, mais il existe d'autres propriétés plus fines; on peut par exemple supposer que les coefficients sont seulement localement lipschitziens en  $x$  et (globalement) à croissance linéaire. Pour plus de renseignements nous renvoyons aux ouvrages cités plus haut.

Nous verrons que nous serons amenés en général à faire des hypothèses plus fortes que les précédentes pour donner des vitesses de convergence de schémas d'approximation numérique. Comme pour les équations différentielles ordinaires, on utilisera en pratique ces résultats de convergence même si les coefficients ont un peu moins de régularité. En toute rigueur il faudrait s'assurer que, pour une équation donnée dont on sait qu'il existe une unique solution forte, ces résultats de convergence sont encore valables en faisant une preuve *à la main*.

### 3.1.3 Modélisation Itô ou Stratonovich ?

Revenons à notre problème de modélisation d'un phénomène bruité abordé en 3.1.1. Si le bruit  $\xi_t$  est très proche d'un bruit blanc, i.e si les fluctuations sont très rapides et pratiquement non corrélées, alors l'équation différentielle stochastique au sens de Itô définie précédemment est un "bon modèle": En pratique on sait par expérience que la solution de l'E.D.S rend bien compte du phénomène bruité.

C'est aussi le cas quand on remplace un modèle discret par un modèle continu. Supposons par exemple que  $X_k$  modélise l'évolution d'une population, et que  $X_k$  vérifie une équation de la forme

$$X_{k+1} = X_k + f(X_k)(t_{k+1} - t_k) + \sigma(X_k)W_k,$$

où  $t_k$  représente une suite de temps et  $W_k$  une suite de gaussiennes indépendantes. Si le pas de temps  $t_{k+1} - t_k$  est suffisamment petit, alors intuitivement l'équation (3.6) est un bon modèle au sens où la solution continue est "proche" du modèle discret. Nous aurons l'occasion au chapitre 4 de mieux formaliser ceci en étudiant la convergence du schéma d'Euler.

Mais dans certaines applications le bruit  $\xi_t$  n'est pas un bruit blanc, mais un processus gaussien physiquement réalisable. Nous renvoyons à Arnold[74] pour des compléments sur le sujet. On peut par exemple utiliser une suite  $\xi_t^n$  vérifiant

$$\mathbf{E}\xi_t^n = 0, \quad \mathbf{E}(\xi_t^n \xi_s^n) = C_n(t - s) \text{ avec } \lim_{n \rightarrow \infty} C_n(t) = \delta(t).$$

Supposons que notre modèle physique soit de la forme

$$\dot{x}_t^n = ax_t^n + bx_t^n \xi_t^n,$$

$n$  étant très grand. On peut penser de premier abord prendre comme modèle l'E.D.S

$$dX_t = aX_t dt + bX_t dW_t. \quad (3.9)$$

Supposons aussi que le bruit soit suffisamment régulier pour que l'on puisse considérer l'équation précédente comme une équation différentielle ordinaire. La solution est alors

$$x_t^n = C \exp(at + b \int_0^t \xi_s^n ds).$$

Introduisons maintenant la diffusion  $X_t = C \exp(at + bW_t)$ . Quand  $n$  tend vers l'infini, la solution limite dans  $L^2$ , notée  $x_t$ , a la même loi que  $X_t$ , qui vérifie d'après la formule d'Itô l'E.D.S

$$dX_t = (a + \frac{1}{2}b^2)X_t + bX_t dW_t.$$

Dans ce cas l'E.D.S au sens d'Itô ne rend pas bien compte du phénomène physique car un terme  $1/2 b^2$  apparait qui ne correspond pas à l'équation (3.9). Expliquons maintenant pourquoi un tel phénomène est possible et comment écrire le bon modèle dans ce cas.

Nous avons rappelé en (3.8) la définition de l'intégrale stochastique au sens de Itô. Il se trouve que d'autres définitions étaient possibles. On aurait pu par exemple définir l'intégrale par

$$\int_0^T f(s, X_s) \circ dW_s = \lim_{\delta_n \rightarrow 0} \sum_k f(t_k^n, \frac{X_{t_k^n} + X_{t_{k+1}^n}}{2})(W_{t_{k+1}^n} - W_{t_k^n}). \quad (3.10)$$

On définit ainsi l'intégrale *au sens de Stratonovich* (voir par exemple Arnold[74] pour compléments). Cette intégrale vérifie les règles du calcul différentiel ordinaire, contrairement à la formule d'Itô, mais par contre ne conserve pas la propriété de martingale, ce qui explique qu'elle soit moins utilisée que l'intégrale au sens de Itô. Rappelons simplement le lien entre les deux intégrales, nous renvoyons à la littérature déjà citée pour plus d'informations.

Les deux représentations suivantes sont équivalentes:

$$dX_t = f(t, X_t)dt + \sigma(t, X_t) \circ dW_t,$$

$$dX_t = (f(t, X_t) + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^n (\frac{\partial \sigma_{.,j}}{\partial x_k} \sigma_{k,j})(t, X_t))dt + \sigma(t, X_t)dW_t,$$

où  $\frac{\partial \sigma_{.,j}}{\partial x_k}$  représente le vecteur colonne de terme général  $\frac{\partial \sigma_{i,j}}{\partial x_k}$ .

On peut maintenant revenir au problème précédent en disant que la solution de l'E.D.S (3.9), *interprétée au sens Stratonovich*, est un bon modèle de l'équation précédente pour  $x_t^n$  quand  $n$  est très grand. Plus généralement un tel phénomène peut être expliqué par la propriété suivante due à Wong et Zakai puis généralisée par la suite (voir par exemple Arnold[74]):

Soit  $X_t^n$  la solution de

$$X_t^n = X_0^n + \int_0^t f(s, X_s^n) ds + \int_0^t \sigma(s, X_s^n) dW_s^n,$$

où le processus  $W_t^n$  est suffisamment régulier et converge pour la norme de la convergence uniforme vers le brownien quand  $n$  tend vers l'infini. Alors la solution  $X_t^n$  de l'équation différentielle ordinaire précédente converge uniformément p.s vers la solution de l'E.D.S correspondante, interprétée au sens Stratonovich.

Il existe de nombreuses généralisations de cette propriété, nous retiendrons simplement *la règle* suivante:

- Si  $\xi_t$  est un "presque" un bruit blanc ou bien si l'on modélise un phénomène discret, alors la "bonne" modélisation est donnée par l'E.D.S interprétée au sens de Itô. On dit que l'on a évalué l'équation différentielle stochastique après avoir fait l'approximation bruit blanc.
- Si le bruit est seulement physiquement réalisable et suffisamment régulier, alors la "bonne" modélisation est donnée par l'E.D.S interprétée au sens de Stratonovich. On dit que l'on a évalué l'équation différentielle ordinaire avant de faire l'approximation bruit blanc.

Il est parfois difficile de savoir laquelle de ces deux modélisations est la bonne, cette question a été discutée par de nombreux auteurs et notamment Stratonovich. Il en est de même pour savoir quel bruit est un bon modèle (bruit blanc ou coloré, gaussien ou poissonnien..) ou encore quelle régularité doit avoir le bruit pour qu'il y ait convergence vers la solution stratonovich, le cas de la dimension  $d$  étant assez délicat. Nous renvoyons à Arnold[74] et aux références incluses dans cet ouvrage.

Pour clore cette introduction disons que le champ d'application des E.D.S est très large, et que l'on aurait tort de croire qu'il se limite à la modélisation du mouvement d'une particule élémentaire. On peut citer entre autre:

- En théorie du filtrage où l'on cherche à observer "au mieux" un signal perturbé par un bruit. Voir pour introduction Arnold[74] ou encore Oksendal[85], et par exemple Kallianpur[80] pour des compléments.
- En théorie du contrôle optimal où l'on cherche à minimiser le coût d'un contrôle en contexte aléatoire. Voir les deux références précédentes comme introduction et Krylov[80] pour compléments.
- En mathématiques financières: Le cours d'une option est modélisé par une E.D.S, permettant ainsi le calcul de la prime de l'option. On renvoie à l'ouvrage récent de Lamberton-Lapeyre[91] et aussi à Karatzas-Shreve[88].
- En mécanique aléatoire: Par exemple pour étudier des phénomènes de houle en liaison avec la théorie de la Markovianisation. Voir Kree-Soize[83].
- En hydrologie où le brownien intervient dans des problèmes de turbulence. Voir par exemple Unny[84].



- En biologie comme modèle limite d'évolution d'une population. Voir par exemple l'ouvrage de Gard[88].
- Et encore en physique théorique, en chimie des réactions... (Van Kampen[81]).

Cette liste n'étant pas exhaustive.

## 3.2 Qu'est ce qu'un schéma de discrétisation ?

### 3.2.1 Pourquoi discrétiser les E.D.S. ?

On ne sait pas en général résoudre explicitement les E.D.S sauf cas particulier comme par exemple le cas où les coefficients  $b$  et  $\sigma$  sont linéaires (voir Arnold[74] ou Gard[88]), alors que faire pour obtenir des informations sur la solution ?

Il existe des méthodes permettant de prévoir à priori le comportement *qualitatif* de la solution. Précisons en quelques mots de quoi il s'agit.

- Stabilité d'une solution: L'idée est de prévoir si deux solutions distinctes, partant de points proches au temps zéro, restent proches au cours du temps. Nous renvoyons à Arnold[74] pour une introduction et aux références incluses dans cet ouvrage.
- Non explosion d'une solution: Si les coefficients de l'équation ne sont pas à croissance linéaire alors en général la solution n'est définie que jusqu'à un temps aléatoire où elle "explose". Il existe des méthodes permettant de prévoir à priori si ce temps d'explosion est fini ou infini. Voir Arnold[74] et références incluses dans cet ouvrage.
- Comportement asymptotique et propriétés de récurrence. On peut prévoir ces propriétés et obtenir certaines estimations à priori. Le fait que la diffusion puisse être dégénérée joue un rôle essentiel dans ces propriétés. On renvoie à l'ouvrage classique de Friedman[75][76].
- Ergodicité de la solution: Dans certains cas la solution possède une unique loi invariante, et ceci peut se prévoir à partir des coefficients  $b$  et  $\sigma$ .

Cependant, pour obtenir des informations *quantitatives* sur les diffusions, il faut recourir à des méthodes numériques. Différentes approches sont possibles comme:

- Considérer les diffusions comme des processus de Markov et faire une approximation en temps et en espace par des chaînes de Markov.
- Se ramener à une équation différentielle ordinaire en approchant le brownien par des fonctions plus régulières.
- Ecrire directement des schémas de discrétisation de (3.6) comme on le fait pour les équations différentielles ordinaires.

La première méthode, bien que très générale, ne tient pas compte de la structure particulière de (3.6); l'utilisation de matrices de transition peut s'avérer lourde, surtout si la dimension  $n$  devient grande. On renvoie à l'ouvrage classique de Kushner[74] sur le sujet.

La deuxième, très importante d'un point de vue théorique, est beaucoup plus délicate d'utilisation en pratique; on est en effet ramené à une équation différentielle ordinaire que l'on ne sait pas non plus résoudre explicitement. On renvoie par exemple à Wong-Zakai[65] et Ikeda-Watanabe[81] pour des compléments.

C'est la troisième méthode que nous allons étudier en détail tout au long des chapitres suivants.

Il existe sur le sujet une cinquantaine d'articles et communications diverses, ainsi que deux ouvrages, Gard[88] qui consacre un chapitre à la discrétisation des E.D.S et un deuxième en préparation de Kloeden-Platen[91]. Il existe aussi un générateur automatique de programmes Fortran de schémas de discrétisation d'E.D.S (voir Talay[90a] ou Leblong-Talay[86]).

On peut faire remonter le début de la discrétisation des E.D.S à Maruyama en 1955 qui introduit l'analogue du schéma d'Euler pour les équations différentielles ordinaires, et en prouve la convergence dans un cadre stochastique. Mais c'est Milshtein en 1973 qui donne véritablement le coup d'envoi à ce domaine et nous aurons l'occasion de reparler de ses travaux.

C'est donc un domaine de recherche récent, et par conséquent en pleine évolution, où beaucoup de questions ne sont pas encore résolues, ni même encore bien posées... S'il existe en effet de nombreux travaux théoriques, ces derniers ne sont parfois pas pertinents dans les applications, par exemple parce que certains schémas de discrétisation performants d'un point de vue théorique se révèlent très difficilement simulables sur ordinateur.

On peut distinguer grossièrement trois classes de travaux qui ont pour thème:

- L'étude de schémas de discrétisation par la recherche de théorèmes mesurant des *vitesse de convergence théoriques*, suivant différents critères.
- La simulation numérique de ces schémas ainsi que des vérifications pratiques de vitesse de convergence sur des exemples *ad hoc*.
- L'utilisation *effective* de ces méthodes de discrétisation pour résoudre des problèmes concrets.

Il faut admettre, et cela résume tout le chemin qu'il reste encore à parcourir en comparaison de ce qui existe sur les équations différentielles ordinaires, que les publications existantes concernent surtout le premier point, un peu le deuxième, et très peu le troisième !

### 3.2.2 Définitions et notations.

On considérera une subdivision, à priori régulière, de l'intervalle de temps  $[0, T]$ , sur lequel on veut approximer la diffusion, notée

$$t_0 = 0 < t_1 < \dots < t_N = T,$$

de pas de discrétisation

$$t_{k+1} - t_k = \Delta t = \frac{T}{N}.$$

Un schéma de discrétisation sera noté  $\{\bar{X}_{t_k}^N\}$  et sera en général défini de façon itérative par une formule du type:

$$\bar{X}_{t_{k+1}}^N = f(t_k, \bar{X}_{t_k}^N, t_{k+1} - t_k, W_{t_{k+1}} - W_{t_k}) , \quad (3.11)$$

où la fonction  $f$  est un polynome simple dépendant de  $b$ ,  $\sigma$  et de leurs dérivées.

Un des intérêts d'une telle représentation est qu'il suffit de savoir simuler les variables aléatoires  $W_{t_{k+1}} - W_{t_k}$ , i.e des gaussiennes indépendantes de variance  $\Delta t$ , pour simuler le schéma de discrétisation. Ces schémas très simples ont cependant leurs limites comme nous le verrons au chapitre 7.

Il faut ensuite définir le critère de convergence le plus approprié au problème étudié. Il existe en effet différents critères de convergence, qui correspondent à des motivations différentes. Ces critères sont principalement de trois types:

- **Convergence quadratique:** On impose au schéma de converger en moyenne quadratique quand le nombre de pas de discrétisation  $N$  tend vers l'infini:

$$\mathbf{E}|X_T - \bar{X}_T^N|^2 \xrightarrow{N} 0.$$

- **Convergence p.s ou trajectorielle:** On impose au schéma de converger p.s vers la solution quand  $N$  tend vers l'infini,

$$\sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N| \xrightarrow{N} 0 \text{ p.s.}$$

- **Convergence en loi ou encore de type Monte Carlo:** Pour une certaine classe de fonction  $f$ , par exemple à croissance polynomiale, on impose au schéma de converger en loi vers la solution,

$$|\mathbf{E}f(X_T) - \mathbf{E}f(\bar{X}_T^N)| \xrightarrow{N} 0 .$$

Disons pour fixer les idées, et en caricaturant quelque peu la réalité, que la situation est radicalement différente de celle des équations différentielles ordinaires:

- On ne peut pas écrire des schémas de discrétisation qui soient à la fois aisément simulables et qui possèdent un ordre arbitrairement grand de convergence.
- Suivant le critère de convergence choisi, les schémas de discrétisation les plus performants ne sont pas du tout les mêmes.
- De nombreux problèmes, comme l'étude de la stabilité des schémas, sont beaucoup plus difficile à résoudre.

Nous reviendrons sur ces points aux chapitres 7 et 8.

### 3.3 Plan des chapitres sur la simulation des diffusions.

Nous commençons par étudier en détail le *schéma d'Euler* au chapitre 4, qui est l'analogue du schéma classique d'Euler pour les équations différentielles ordinaires. Le schéma d'Euler, par sa simplicité, permet une introduction aisée au domaine de la discrétisation des E.D.S. Nous donnons des résultats nouveaux de convergence  $L^P$  et  $p.s$  du schéma, en généralisant le résultat de Maruyama[55] et en précisant les vitesses de convergence. Nous présentons aussi des résultats pour le critère de convergence en loi.

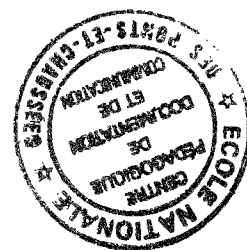
Nous étudions ensuite au chapitre 5 le *schéma de Milstein* qui permet, dans certains cas, d'améliorer la vitesse de convergence du schéma d'Euler. Nous donnons des résultats nouveaux sur les vitesses de convergence  $L^P$  et  $p.s$  en améliorant le résultat de Milstein[74] et partiellement Talay[83b].

Puis nous abordons au chapitre 6 une autre classe de schémas où le pas de discrétisation n'est plus fixe comme précédemment, mais variable, afin par exemple de *contrôler spatialement* le brownien ou même la diffusion elle-même. Deux schémas de discrétisation sont introduits, dont on précise les vitesses de convergence et dont on donne un algorithme de simulation. Ces résultats sont une généralisation de Bichteler[81] et Newton[90a]. Une application à l'étude des diffusions qui explosent est donnée.

Au chapitre 7 nous dressons un *panorama des travaux existants sur la discrétisation des E.D.S.* Il nous a paru en effet nécessaire de faire une présentation synthétique permettant à la fois une utilisation immédiate en simulation, sans besoin de se référer à de nombreux articles, mais aussi de disposer d'une vision d'ensemble permettant, nous l'espérons, de mieux comprendre les problèmes posés par la discrétisation des E.D.S. Loin d'être exhaustif, cette synthèse présente quelques morceaux choisis qui nous semblent pertinents pour les applications et il y a donc de nombreuses omissions de travaux importants. Toutefois nous renvoyons à de nombreuses références.

Précisons que cette partie peut, dans une large mesure, être lue indépendamment des parties précédentes.

Enfin nous abordons au chapitre 8 *quelques problèmes ouverts* qui semblent essentiels pour les applications et qui n'ont, pour le moment, que des réponses très partielles.





## Chapitre 4

### Etude du schéma d'Euler.



## 4.1 Introduction.

Soit  $\{X_t\}$  une diffusion solution de l'équation différentielle stochastique de type Itô:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad t \in [0, T], \quad (4.1)$$

avec  $b: \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ ,  $\sigma: \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^{n \times d}$  et  $(W_t)$  un brownien standard de  $\mathbf{R}^d$ .

Il n'est en général pas possible de résoudre explicitement (4.1) et des méthodes numériques sont nécessaires pour obtenir des informations quantitatives sur la solution comme nous l'avons vu en 3.2.

Historiquement le premier schéma de discrétisation a été proposé par Maruyama[55] comme une généralisation du schéma d'Euler pour les équations différentielles ordinaires.

Si on partitionne l'intervalle  $[0, T]$  en une subdivision régulière  $t_0 = 0 < t_1 < \dots < t_N$  de pas  $\Delta t = T/N$ , alors le schéma d'Euler-Maruyama est défini par  $\bar{X}_0^N = X_0$  et

$$\bar{X}_{t_{k+1}}^N = \bar{X}_{t_k}^N + \sigma(t_k, \bar{X}_{t_k}^N)(W_{t_{k+1}} - W_{t_k}) + b(t_k, \bar{X}_{t_k}^N)(t_{k+1} - t_k).$$

Maruyama[55] a montré la convergence quadratique de ce schéma vers la solution de (4.1) au sens suivant:

$$\exists C > 0, \forall k = 0..N, \mathbf{E}|X_{t_k} - \bar{X}_{t_k}^N|^2 \leq C \Delta t. \quad (4.2)$$

L'intérêt d'un tel schéma est qu'il est très facile de le simuler car la seule v.a à générer, i.e  $W_{t_{k+1}} - W_{t_k}$ , est une gaussienne de variance  $t_{k+1} - t_k$ .

Ce schéma a été depuis abondamment étudié et, comme nous le verrons, il existe d'autres schémas de discrétisation qui permettent souvent une convergence à un ordre plus élevé, par exemple en  $(\Delta t)^2$ . Notre objectif est de donner, sur ce schéma très simple, les résultats connus suivant différents critères de convergence possibles (voir 3.2), en donnant des démonstrations élémentaires et complètes.

Ce schéma, par sa simplicité, peut être considéré comme un bon outil pédagogique d'introduction aux méthodes numériques de discrétisation des E.D.S. De plus, pour une diffusion quelconque dans  $\mathbf{R}^n$ , ce schéma est dans certains cas aussi performant que les autres comme nous le verrons et il mérite d'être considéré même en vue des applications.

Les résultats nouveaux de ce chapitre sont la convergence  $L^p$  et  $p.s$  du schéma d'Euler. Après avoir prouvé ces résultats nous avons découvert que des résultats similaires apparaissent dans Kanagawa[88] pour la convergence  $L^p$  et Newton[90b] pour la convergence  $p.s$ . Toutefois notre résultat est plus général que les deux précédents comme nous allons le voir.

**Précisons les notations:** On considère donc une subdivision régulière  $0 = t_0 < t_1 < \dots < t_N$  de pas constant  $\Delta t = \frac{T}{N}$ . On définit le schéma d'Euler par  $\bar{X}_0^N = X_0$  et

$$\forall t \in [t_k, t_{k+1}], \bar{X}_t^N = \bar{X}_{t_k}^N + \sigma(t_k, \bar{X}_{t_k}^N)(W_t - W_{t_k}) + b(t_k, \bar{X}_{t_k}^N)(t - t_k). \quad (4.3)$$

On peut remarquer que l'on retrouve bien le schéma d'Euler classique aux points  $t_k$  de la subdivision et que tous les résultats qui suivront seront valables pour le schéma discret défini par Maruyama.

L'erreur d'approximation au temps  $t$  est notée  $\varepsilon_t = X_t - \bar{X}_t^N$ .

Enfin  $|x|$  représente la norme euclidienne sur  $\mathbf{R}^n$  et on note  $|\dot{M}|^2 = \text{trace}(M.M')$  pour  $M \in \mathbf{R}^{n \times d}$ .



**Précisons les hypothèses:** On prend des hypothèses assurant l'existence et l'unicité de la solution de l'E.D.S que nous conserverons tout au long de ce chapitre 4.

(H<sub>1</sub>)  $b$  et  $\sigma$  sont lipschitziennes en  $x$ :

$$\exists K_1 > 0, \forall t \in [0, T], \forall x, y \in \mathbf{R}^n :$$

$$|\sigma(t, x) - \sigma(t, y)| + |b(t, x) - b(t, y)| \leq K_1 |x - y| .$$

(H<sub>2</sub>)  $b$  et  $\sigma$  sont à croissance linéaire, localement uniformément en  $t$ :

$$\exists K_2 > 0, \forall t \in [0, T], \forall x \in \mathbf{R}^n :$$

$$|\sigma(t, x)|^2 + |b(t, x)|^2 \leq K_2^2 (1 + |x|^2) .$$

(H<sub>3</sub>) Il existe  $p \geq 1$  tel que  $\mathbf{E}|X_0|^{2p} < +\infty$ .

(H<sub>4</sub>)  $b$  et  $\sigma$  sont höldériennes en  $t$ :

$$\exists K_4 > 0, \exists \beta > 0, \forall s, t \in [0, T], \forall x \in \mathbf{R}^n :$$

$$|\sigma(t, x) - \sigma(s, x)| + |b(t, x) - b(s, x)| \leq K_4 |t - s|^\beta .$$

(H<sub>1</sub>), (H<sub>2</sub>) et (H<sub>3</sub>) garantissent l'existence et l'unicité d'une solution pour l'équation (4.1) qui vérifie de plus pour  $q \leq p$  (Friedman[75]) les deux propriétés suivantes:

$$(P_1) \quad \mathbf{E}|X_t|^{2q} \leq (1 + \mathbf{E}|X_0|^{2q})e^{C_1 t} \quad \forall t \in [0, T] ,$$

et

$$(P_2) \quad \mathbf{E}|X_t - X_s|^{2q} \leq C_2 (1 + \mathbf{E}|X_0|^{2q}) |t - s|^q ,$$

où les constantes  $C_1$  et  $C_2$  ne dépendent que de  $K_2$ ,  $p$  et  $T$ .

## 4.2 Convergence $L^p$ et $p.s$ du schéma

Nous donnons le résultat de convergence  $L^p$  et  $p.s$  du schéma d'Euler. Comme annoncé, un résultat similaire apparaît pour la convergence  $L^p$  dans Kanagawa[88] mais sous l'hypothèse  $b$  et  $\sigma$  bornés, et un résultat similaire apparaît pour la convergence  $p.s$  dans Newton[90b], mais au temps final  $T$  et non uniformément sur  $[0, T]$ . On a la

**Proposition 14.** *Sous les hypothèses (H<sub>1</sub>) à (H<sub>4</sub>), on a*

$$\mathbf{E}(\sup_{t \in [0, T]} |X_t - \bar{X}_t^N|^{2p}) \leq C_3 \Delta t^\gamma \quad \text{avec } \gamma = p \inf(2\beta, 1). \quad (4.4)$$

La constante  $C_3$  ne dépend que de  $K_1, K_2, K_4, T, p$  et  $\mathbf{E}|X_0|^{2p}$ .

Si de plus il existe  $p > 1$  vérifiant  $(H_3)$  et tel que  $\gamma > 1$ , alors il y a convergence p.s quand  $N$  tend vers l'infini:

$$\sup_{t \in [0, T]} |X_t - \bar{X}_t^N| \xrightarrow{p.s} 0. \quad (4.5)$$

Si de plus  $(H_3)$  est vérifiée pour tous  $p$  et si  $(H_4)$  est vérifiée pour  $\beta \geq \frac{1}{2}$  alors:

$$\forall \alpha < \frac{1}{2}, \quad N^\alpha \sup_{t \in [0, T]} |X_t - \bar{X}_t^N| \xrightarrow{p.s} 0. \quad (4.6)$$

**Remarque 7** Dans le cas homogène  $\sigma(t, x) = \sigma(x)$  et  $b(t, x) = b(x)$ ,  $(H_1)$  et  $(H_3)$  suffisent et les mêmes résultats sont valables avec  $\gamma = p$ .

Donnons maintenant une propriété qui montre que la vitesse de convergence p.s du schéma d'Euler n'est pas meilleure que  $\frac{1}{\sqrt{N}}$ . On a la

**Proposition 15.** Dans le cas linéaire en dimension 1,  $b(x) = 0$ ,  $\sigma(x) = x$  et  $X_0 = 1$ , on a

$$\forall \alpha > \frac{1}{2}, \quad \limsup_N N^\alpha |X_T - \bar{X}_T^N| = +\infty \text{ p.s.} \quad (4.7)$$

**Remarque 8** Dans l'exemple précédent il est possible d'obtenir une convergence en loi quand  $\alpha = \frac{1}{2}$ . Cette propriété est en fait généralisable en faisant quelques hypothèses sur  $\sigma$  (voir Newton[90b]).

En pratique si l'on veut disposer d'une approximation de la diffusion en tous les points de l'intervalle  $[0, T]$ , on peut effectuer une interpolation linéaire entre les points de la subdivision. Nous allons voir que ce schéma d'Euler interpolé linéairement possède presque la même vitesse de convergence que le schéma Euler classique.

Pour cela introduisons le schéma d'Euler Interpolé Linéairement (E.I.L.) défini par  $\bar{X}_0^N = X_0$  et:

$$\begin{aligned} \bar{X}_{t_{k+1}}^N &= \bar{X}_{t_k}^N + \sigma(t_k, \bar{X}_{t_k}^N)(W_{t_{k+1}} - W_{t_k}) + b(t_k, \bar{X}_{t_k}^N)(t_{k+1} - t_k) \\ \forall t \in [t_k, t_{k+1}[ , \bar{X}_t^N &= \bar{X}_{t_k}^N + \frac{t - t_k}{t_{k+1} - t_k}(\bar{X}_{t_{k+1}}^N - \bar{X}_{t_k}^N). \end{aligned}$$

On aurait aussi pu prendre un schéma constant entre les points de la subdivision. On a alors la proposition suivante donnée dans le cas où  $\gamma \geq 1/2$  et où  $(H_3)$  est vérifiée pour tous  $p$  pour simplifier l'énoncé.

**Proposition 16.** Sous les hypothèses  $(H_1)$ ,  $(H_2)$ ,  $(H_3) \forall p$ ,  $(H_4)$  avec  $\gamma \geq 1/2$ , le schéma E.I.L. vérifie:

$$\mathbb{E} \left( \sup_{t \in [0, T]} |X_t - \bar{X}_t^N|^{2p} \right) \leq C_4 (\Delta t)^p (-\text{Log}(\Delta t))^p. \quad (4.8)$$

La constante  $C_4$  ne dépend que de  $K_1, K_2, K_4, T, p$  et  $\mathbb{E}|X_0|^{2p}$ . De plus il y a convergence trajectorielle à la même vitesse que le schéma d'Euler de type diffusion:

$$\forall \alpha < \frac{1}{2}, \quad N^\alpha \sup_{t \in [0, T]} |X_t - \bar{X}_t^N| \xrightarrow{p.s} 0. \quad (4.9)$$

Démontrons ces propriétés. Commençons par deux lemmes. On dit qu'un processus  $Y_t$  est  $L^p$ -continu si la fonction de  $t$ ,  $E|Y_t|^p$  est continue. On a alors:

**Lemme 1.** Si  $Y_t = Y_0 + \int_0^t A_s dW_s + \int_0^t B_s ds$  avec  $A_s$  et  $B_s$   $L^{2p}$ -continus, alors  $Y_t$  est  $L^{2p}$ -continu et de plus

$$E|Y_t|^{2p} \leq E|Y_0|^{2p} + C_5 \int_0^t E\{|Y_s|^{2p} + |A_s|^{2p} + |B_s|^{2p}\} ds .$$

Preuve du lemme 1: Commençons par remarquer que, d'après  $(x+y+z)^{2p} \leq C_p(x^{2p}+y^{2p}+z^{2p})$  pour tous  $x, y, z \geq 0$ ,  $Y_t$  est  $L^{2p}$ -continu et  $\int_0^t E|Y_s|^{2p} ds$  est en particulier finie pour tous  $t$ . Afin d'utiliser la formule d'Itô pour obtenir le résultat recherché, nous utilisons un argument de localisation. Soit  $\tau_N = \inf\{t > 0, |Y_t| > N\}$  et  $Y_t^{(N)} = Y_{t \wedge \tau_N}$ . En appliquant la formule d'Itô à  $Y_t^{(N)}$  on obtient,

$$\begin{aligned} |Y_t^{(N)}|^{2p} &= |Y_0|^{2p} + \int_0^t 2p|Y_s^{(N)}|^{2p-2} Y_s^{(N)} A_s dW_s + \int_0^t 2p|Y_s^{(N)}|^{2p-2} Y_s^{(N)} B_s ds \\ &\quad + \int_0^t (p|Y_s^{(N)}|^{2p-2} |A_s|^2 + 2p(p-1)|Y_s^{(N)}|^{2p-4} |Y_s^{(N)} A_s|^2) ds . \end{aligned}$$

L'intégrale stochastique est une vraie martingale et son espérance est donc nulle. Ainsi on obtient:

$$E|Y_t^{(N)}|^{2p} \leq E|Y_0|^{2p} + C \int_0^t E\{|Y_s^{(N)}|^{2p} + |B_s|^2 |Y_s^{(N)}|^{2p-2} + |A_s|^2 |Y_s^{(N)}|^{2p-2}\} ds .$$

Enfin, en utilisant  $x^{2p-2}y^2 \leq Cx^{2p} + Cy^{2p}$  pour tous  $x, y \geq 0$  et le lemme de Fatou quand  $N$  tend vers l'infini, on conclut.

□

On utilisera aussi une version classique du lemme de Borel-Cantelli:

**Lemme 2.** Soit  $(X_n)_n$  une suite de v.a telle que

$$\exists \alpha > 0, \beta > 1, \quad E(|X_n|^\alpha) \leq N^{-\beta}$$

alors  $X_n$  converge p.s vers zero.

Preuve du lemme 2: Si  $x_n = |X_n|^\alpha$ , on a  $\sum_n E(x_n) < +\infty$ . Ainsi la série  $\sum_n x_n$  est p.s convergente, et son terme général converge donc p.s vers zéro.

□

**Preuve de la proposition 14:** . On utilise trois étapes:

**.Première étape:** On montre que

$$\sup_{k \in [0, N]} E(|\varepsilon_{t_k}^N|^{2p}) \leq C(\Delta t)^\gamma .$$

Remarquons que pour  $t \in [t_k, t_{k+1}]$ , on a:

$$\varepsilon_t^N = \varepsilon_{t_k}^N + \int_{t_k}^t (\sigma(s, X_s) - \sigma(t_k, \bar{X}_{t_k}^N)) dW_s + \int_{t_k}^t (b(s, X_s) - b(t_k, \bar{X}_{t_k}^N)) ds .$$

Il est clair par ailleurs que le schéma d'Euler vérifiant l'E.D.S (4.3), il vérifie aussi les propriétés (P1) et (P2).

Posons  $Y_u = \varepsilon_{t_k+u}^N$  pour  $u \in [0, t - t_k]$ . Le lemme 1 appliqué à  $Y_u$ , les intégrants étant  $L^{2p}$ -continus d'après (P1) et (P2), donne pour  $t \leq t_{k+1}$ :

$$\mathbf{E}(|\varepsilon_t^N|^{2p}) \leq \mathbf{E}(|\varepsilon_{t_k}^N|^{2p}) + C \int_{t_k}^t \mathbf{E}\{|\varepsilon_s^N|^{2p} + |\sigma(s, X_s) - \sigma(t_k, \bar{X}_{t_k}^N)|^{2p} + |b(s, X_s) - b(t_k, \bar{X}_{t_k}^N)|^{2p}\} ds.$$

De plus

$$\begin{aligned} |b(s, X_s) - b(t_k, \bar{X}_{t_k}^N)|^{2p} &\leq C|b(s, X_s) - b(t_k, X_s)|^{2p} + C|b(t_k, X_s) - b(t_k, X_{t_k})|^{2p} \\ &\quad + C|b(t_k, X_{t_k}) - b(t_k, \bar{X}_{t_k}^N)|^{2p}. \end{aligned}$$

Ainsi avec  $(H_1)$ ,  $(H_4)$  et (P2), en notant  $\gamma = p \inf(2\beta, 1)$ , on obtient:

$$\mathbf{E}\{|b(s, X_s) - b(t_k, \bar{X}_{t_k}^N)|^{2p}\} \leq C(\Delta t)^\gamma + C\mathbf{E}(|\varepsilon_{t_k}^N|^{2p}),$$

et une inégalité similaire pour  $\sigma$ .

En écrivant  $\varepsilon_s^N = \varepsilon_{t_k}^N + (X_s - X_{t_k}) + (\bar{X}_{t_k}^N - \bar{X}_s^N)$  et avec  $(H_4)$  on a

$$\mathbf{E}|X_s - X_{t_k}|^{2p} \leq C(\Delta t)^p,$$

et une expression similaire pour  $\bar{X}_t^N$ . Ainsi il vient

$$\mathbf{E}(|\varepsilon_{t_{k+1}}^N|^{2p}) \leq \mathbf{E}|\varepsilon_{t_k}^N|^{2p} (1 + C\Delta t) + C(\Delta t)^{\gamma+1}.$$

On conclut alors par l'argument suivant :

Si  $x_{n+1} \leq a x_n + b$  alors  $x_n \leq a^n x_0 + n e^{n(a-1)} b$  pour  $a \geq 1$ .

Ainsi on obtient,

$$\mathbf{E}|\varepsilon_{t_{k+1}}^N|^{2p} \leq N e^{NC\Delta t} C(\Delta t)^{\gamma+1},$$

puis

$$\mathbf{E}|\varepsilon_{t_{k+1}}^N|^{2p} \leq C(\Delta t)^\gamma.$$

□

**.Deuxième étape:** On montre que

$$\mathbf{E}(\sup_{t \in [0, T]} |\varepsilon_t^N|^{2p}) \leq C(\Delta t)^\gamma.$$

En utilisant

$$\begin{aligned} \varepsilon_t^N &= \int_0^t \sum_{k=0}^{N-1} (\sigma(s, X_s) - \sigma(t_k, \bar{X}_{t_k}^N) \mathbf{1}_{]t_k, t_{k+1}]}(s)) dW_s \\ &\quad + \int_0^t \sum_{k=0}^{N-1} (b(s, X_s) - b(t_k, \bar{X}_{t_k}^N) \mathbf{1}_{]t_k, t_{k+1}]}(s)) ds, \end{aligned}$$

et l'inégalité de Burkholder-Davis-Gundy, on a

$$\begin{aligned} \mathbf{E}(\sup_{t \in [0, T]} |\varepsilon_t^N|^{2p}) &\leq C \mathbf{E}\left\{\int_0^T \sum_{k=0}^{N-1} |\sigma(s, X_s) - \sigma(t_k, \bar{X}_{t_k}^N)|^{2p} \mathbf{1}_{[t_k, t_{k+1}]}(s) ds\right\} \\ &+ C \mathbf{E}\left\{\int_0^T \sum_{k=0}^{N-1} |b(s, X_s) - b(t_k, \bar{X}_{t_k}^N)|^{2p} \mathbf{1}_{[t_k, t_{k+1}]}(s) ds\right\}. \end{aligned}$$

Par l'étape 1 on obtient

$$\mathbf{E}|\sigma(s, X_s) - \sigma(t_k, \bar{X}_{t_k}^N)|^{2p} \leq C(\Delta t)^\gamma$$

et une expression similaire pour  $b$ . On en déduit donc:

$$\mathbf{E}(\sup_{t \in [0, T]} |\varepsilon_t^N|^{2p}) \leq C(\Delta t)^\gamma.$$

**Troisième étape:** On montre les résultats de convergence p.s. Si  $\gamma > 1$ , le lemme 2 donne

$$\lim_N (\sup_{t \in [0, T]} |\varepsilon_t^N|) = 0 \text{ p.s.}$$

De plus si  $(H_3)$  est vérifiée pour tous  $p$  et si  $\beta \geq \frac{1}{2}$ , alors pour  $\alpha < 1/2$ , on peut choisir  $p$  tel que  $p(1 - 2\alpha) > 1$ . Le lemme 2 donne alors le résultat.

□

**Preuve de la proposition 15:**

On donne la preuve pour  $T = 1$  et on suppose que  $\alpha < \frac{2}{3}$  (Le cas  $\alpha \geq \frac{2}{3}$  s'en déduit immédiatement).

On a

$$\bar{X}_1 = \prod_{i=1}^N (1 + \Delta W_i) \text{ avec } \Delta W_i = W_{t_i} - W_{t_{i-1}}$$

et  $X_1 = e^{W_1 - \frac{1}{2}}$ . Remarquons tout d'abord que, puisque  $W_t$  est uniformément continu sur  $[0, 1]$ , on a

$$\forall \omega \in \Omega, \exists N_\omega, \forall N \geq N_\omega, \forall i = 1, N, |\Delta W_i| < \frac{1}{2}. \quad (4.10)$$

Puis pour  $N \geq N_\omega$  on obtient

$$\text{Log } \bar{X}_1 = W_1 - \frac{1}{2} \sum_{i=1}^N (\Delta W_i)^2 + O\left(\sum_{i=1}^N (\Delta W_i)^3\right),$$

et ainsi

$$\text{Log } \frac{\bar{X}_1}{X_1} = T_1 + O(T_2),$$

avec

$$T_1 = \frac{1}{2} \left(1 - \sum_{i=1}^N (\Delta W_i)^2\right) \text{ et } T_2 = \sum_{i=1}^N (\Delta W_i)^3.$$

Remarquons que

$$\mathbf{E}(T_1^6) = \frac{C}{N^3} \text{ et } \mathbf{E}(T_2^4) = \frac{C}{N^4},$$

puisque si l'on note  $g_i = \sqrt{N} \Delta W_i$  on a

$$T_1 = \frac{1}{2N} \sum_{i=1}^N (1 - g_i^2).$$

De  $\mathbf{E}(|1 - g_i^2|^p) = C_p$  on conclut facilement. L'argument est similaire pour  $T_2$ .

Par le lemme 2 on obtient

$$N^\alpha T_1^2 \xrightarrow{p.s.} 0 \text{ et } N^\alpha T_2 \xrightarrow{p.s.} 0 \quad (4.11)$$

donc

$$\begin{aligned} \frac{\bar{X}_1}{X_1} &= 1 + T_1 + O(T_1^2) + O(T_2), \\ N^\alpha (\bar{X}_1 - X_1) &= X_1 (N^\alpha T_1 + O(N^\alpha T_1^2) + O(N^\alpha T_2)). \end{aligned} \quad (4.12)$$

De plus

$$\limsup_N |N^\alpha T_1| = +\infty \text{ p.s.} \quad (4.13)$$

En effet

$$N^\alpha T_1 = \frac{1}{2} N^{\alpha-\frac{1}{2}} \sum_{i=1}^N \frac{1 - g_i^2}{\sqrt{N}}$$

et le théorème central limite donne  $|N^\alpha T_1| \xrightarrow{(P)} +\infty$  et montre ainsi (4.13). Puis avec (4.12), (4.11) et (4.13) on obtient le résultat.

□

**Preuve de la proposition 16:** Cette démonstration s'inspire de Kanagawa[88] en simplifiant toutefois sa preuve.

Posons  $\epsilon_t = \bar{X}_t - \tilde{X}_t$ , on a omis les  $N$  pour alléger l'écriture et noté  $\delta$  à la place de  $\Delta t$ . On a pour  $t \in [t_k, t_{k+1}]$ ,

$$\epsilon_t = \sigma(\bar{X}_{t_k})(W_t - W_{t_k} - \frac{t - t_k}{\delta}(W_{t_{k+1}} - W_{t_k})) = \sigma(\bar{X}_{t_k})A_t^k.$$

De

$$|A_t^k| \leq |W_t - W_{t_k}| + |W_t - W_{t_{k+1}}|$$

et

$$\mathbf{E} \left| \sup_{k \in [0, N]} \sigma(\bar{X}_{t_k}) \right|^{4p} < +\infty,$$

il vient

$$\mathbf{E} \left( \sup_{t \in [0, T]} |\epsilon_t|^{2p} \right) \leq C \left\{ \mathbf{E} \left( \sup_{k \in [0, N]} \sup_{t \in [t_k, t_{k+1}]} |A_t^k|^{4p} \right) \right\}^{1/2}.$$

Puis

$$\mathbf{E}(\sup_{t \in [0, T]} |\epsilon_t|^{2p}) \leq C \{ \mathbf{E}(\sup_{k \in [0, N]} \sup_{t \in [0, \delta]} |B_t^k|^{4p}) \}^{1/2} \quad (4.14)$$

où les  $B_t^k$  sont N browniens standards indépendants. Si on note

$$M_\delta = \max_{k \in [0, N]} \sup_{t \in [0, \delta]} |B_t^k|,$$

alors nous allons montrer que

$$(\mathbf{E} M_\delta^{4p})^{1/2} \leq C \delta^p (-\log \delta)^p, \quad (4.15)$$

ce qui achèvera la preuve. Pour cela on écrit que

$$\mathbf{E}(M_\delta^{4p}) \leq \mathbf{E}(M_\delta^{4p} 1_{\{M_\delta > 2\sqrt{\delta}(-\log \delta)^{1/2}\}}) + C \delta^{2p} (-\log \delta)^{2p}.$$

Notons alors

$$I = \mathbf{E}(M_\delta^{4p} 1_{\{M_\delta > 2\sqrt{\delta}(-\log \delta)^{1/2}\}}).$$

On a

$$I \leq \sum_{k=1}^N \mathbf{E}(\sup_{t \in [0, \delta]} |B_t^k|^{4p} 1_{\{\sup_{t \in [0, \delta]} |B_t^k| \geq 2\sqrt{\delta}(-\log \delta)^{1/2}\}}).$$

Puis grâce au principe de réflexion avec

$$g(x) = 2 \frac{e^{-x^2/2\delta}}{\sqrt{2\pi\delta}},$$

on obtient

$$I \leq 2 \sum_{k=1}^N \mathbf{E}(|\sup_{t \in [0, \delta]} B_t^k|^{4p} 1_{\{\sup_{t \in [0, \delta]} B_t^k \geq 2\sqrt{\delta}(-\log \delta)^{1/2}\}})$$

$$I \leq 2N \int_{x \geq 2\sqrt{\delta}(-\log \delta)^{1/2}} x^{4p} g(x) dx.$$

$$I \leq C N \delta^{2p} \int_{u \geq 2(-\log \delta)^{1/2}} u^{4p} \exp(-\frac{u^2}{2}) du.$$

Par récurrence sur p on majore facilement l'intégrale précédente et on obtient

$$I \leq C_p \delta^{2p+1} (-\log \delta)^{2p},$$

ce qui donne (4.15). En utilisant (4.15) et (4.14) on conclut.

Pour la convergence p.s le raisonnement est le même qu'en proposition 14.

□

### 4.3 Résultats numériques

On sait par la proposition 14 que le schéma d'Euler converge *p.s* à la vitesse de  $N^{\frac{1}{2}-\epsilon}$  pour  $\epsilon > 0$ , mais quel choix de  $N$  faut-il faire en pratique pour obtenir une erreur inférieure à un réel fixée ?

Remarquons tout d'abord que pour observer la convergence *p.s* du schéma d'Euler vers une trajectoire de la diffusion, il faut utiliser un raffinement de la trajectoire brownienne dans la simulation. Un tel raffinement peut se faire en utilisant une simulation récursive du mouvement brownien, i.e en construisant des ponts browniens comme exposé en 1.2.

On présente des résultats numériques sur deux exemples. Le premier est une diffusion linéaire en dimension 1, et le deuxième est un système bidimensionnel pour lequel on ne connaissait à priori pas de schémas, facilement simulables, convergeant trajectoriellement (voir 7.2).

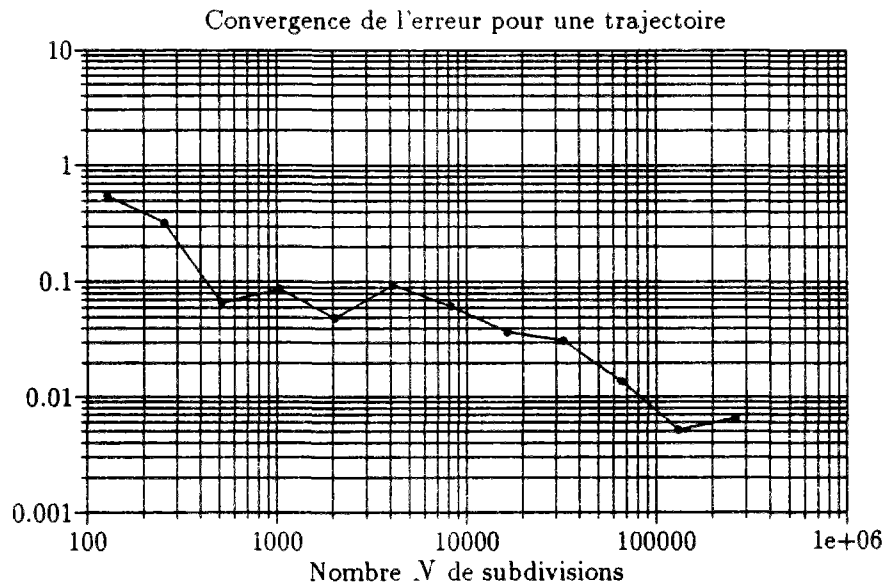
**Exemple 1:** La diffusion  $X_t = e^{W_t}$  avec  $X_0 = 1$  est solution de l'E.D.S:

$$dX_t = \frac{1}{2}X_t dt + X_t dW_t, \quad t \in [0, 1].$$

Afin d'estimer l'erreur d'approximation sur une trajectoire on procède de la façon suivante. On effectue une simulation récursive du brownien et on mesure à chaque étape l'erreur d'approximation

$$e_N = \sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N|$$

où  $N$  représente le nombre de points de la subdivision.





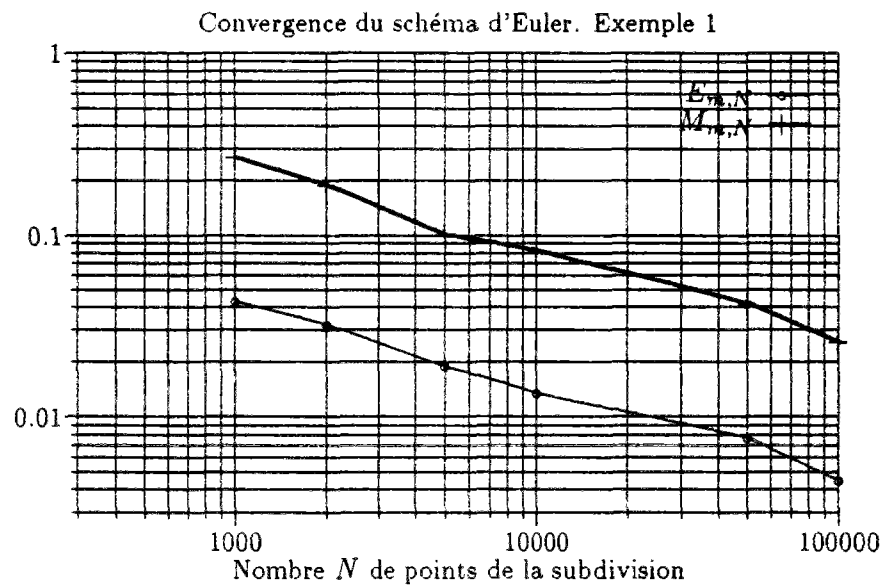
Si  $(e_N^1, \dots, e_N^m)$  sont  $m$  réalisations indépendantes du schéma d'Euler, on estime

$$E_{m,N} = \frac{e_N^1 + \dots + e_N^m}{m}$$

et

$$M_{m,N} = \text{Max}(e_N^1, \dots, e_N^m)$$

On obtient avec  $m = 1000$  la courbe suivante



On constate bien que la vitesse de convergence est d'environ  $N^{-\frac{1}{2}}$ .

**Exemple 2:** Cet exemple a déjà été étudié par McShane[74], Clark-Cameron[80], Rumelin[82] et Pardoux-Talay[85].

Soit  $W^1$  et  $W^2$  deux browniens indépendants. La diffusion solution de

$$X_t^1 = W_t^1$$

$$dX_t^2 = X_t^1 dW_t^2 \text{ et } X_0^2 = 0$$

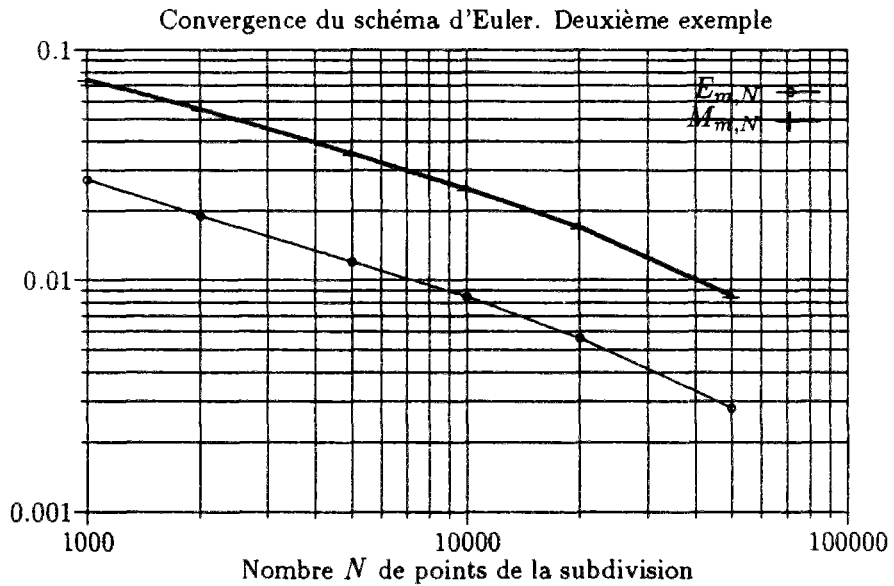
ne vérifie pas la condition classique de commutativité (dont nous reparlerons en 7.2)

$$(\nabla_x \sigma_i)(\sigma_j) = (\nabla_x \sigma_j)(\sigma_i) .$$

Dans ce cas on ne sais pas si le schéma de Milshtein converge trajectoriellement (Talay[83b] et 7.2).

Ici il n'est justement (voir 7.2) pas possible de simuler une trajectoire de la diffusion en utilisant les valeurs  $\{W_{t_k}, t_k \in [0, T]\}$ .

On considère donc que l'erreur d'approximation est  $\sup_{k \in [0, N]} |\bar{X}_{t_k}^M - \bar{X}_{t_k}^N|$  avec  $M = 100\,000$ . Avec les mêmes notations que dans l'exemple 1 on obtient :



La vitesse de convergence est encore en  $N^{-\frac{1}{2}}$ .

#### 4.4 Approximation du flot de l'E.D.S

Afin d'approcher le flot de l'E.D.S (4.1), on effectue une discrétisation en temps de pas  $\Delta t$  comme précédemment et aussi une discrétisation en espace pour les valeurs initiales de pas  $\Delta x$ . On note  $(X_t^x)_t$  la solution de (4.1) telle que  $X_0 = x$  et  $(\bar{X}_t^y)_t$  le schéma d'Euler correspondant à la solution  $(X_t^y)_t$ . On peut étudier la convergence  $L^p$  du flot approché vers le flot exact. On a la

**Proposition 17.** *Sous les hypothèses  $(H_1)$  à  $(H_4)$  de 4.1 on a*

$$\mathbf{E} \left( \sup_{t \in [0, T], |x-y| < \Delta x, |x| \leq M} |X_t^x - \bar{X}_t^y|^{2p} \right) \leq C_6 (\Delta t)^\gamma + C_7 (\Delta x)^{2p}$$

avec,  $\gamma = p \inf(2\beta, 1)$ , les constantes  $C_6$  et  $C_7$  ne dépendant que de  $M, T, p, K_1, K_2, K_4$ .

**Preuve de la proposition 17.** On note  $\varepsilon_t^N = X_t^x - \bar{X}_t^y$ . On a encore

$$\varepsilon_s^N = \varepsilon_{t_k}^N + \int_{t_k}^s (\sigma(u, X_u^x) - \sigma(t_k, \bar{X}_{t_k}^y)) dW_u + \int_{t_k}^s (b(u, X_u^x) - b(t_k, \bar{X}_{t_k}^y)) du.$$

En procédant comme à la proposition 1 on obtient

$$\mathbf{E}(|\varepsilon_{t_{k+1}}^N|^{2p}) \leq (1 + C\Delta t)^k \mathbf{E}(|\varepsilon_0^N|^{2p}) + k e^{kC\Delta t} C(\Delta t)^{\gamma+1},$$

ainsi

$$\mathbf{E}(|\varepsilon_{t_{k+1}}^N|^{2p}) \leq C|x - y|^{2p} + C(\Delta t)^\gamma.$$

En utilisant la même représentation de  $\varepsilon_t^N$  sous forme de diffusion comme précédemment, on obtient

$$\mathbf{E} \left( \sup_{t \in [0, T]} |\varepsilon_t^N - \varepsilon_0^N|^{2p} \right) \leq C(\Delta t)^\gamma + C(\Delta x)^{2p},$$

et ainsi

$$\mathbf{E} \left( \sup_{t \in [0, T]} |\varepsilon_t^N|^{2p} \right) \leq C(\Delta t)^\gamma + C(\Delta x)^{2p}.$$

□

**Corollaire 1.** *Sous les mêmes hypothèses qu'à la Proposition 17 et si de plus*

$$\exists p > 1 \text{ tel que } \gamma > 1 \text{ et } \exists C > 0 \text{ tel que } \Delta x \leq C\sqrt{\Delta t},$$

alors il y a convergence p.s:

$$\sup_{t \in [0, T], |x-y| < \Delta x, |x| \leq M} |X_t^x - \bar{X}_t^y| \xrightarrow{p.s} 0$$

**Preuve:** L'argument est similaire à celui utilisé dans la proposition 14.

□

On donne maintenant une propriété de continuité du flot discrétisé:

**Proposition 18.** *Sous les hypothèses  $(H_1)$  à  $(H_4)$  de 4.1 on a:*

$$\mathbf{E} \left( \sup_{t \in [0, T], |x| \leq M} |\bar{X}_t^x - \bar{X}_t^y|^{2p} \right) \leq C |x - y|^{2p}.$$

*La constante  $C$  ne dépendant que de  $T, p, K_1$ .*

**Preuve de la proposition 18:** Notons  $\varepsilon_t^N = \bar{X}_t^x - \bar{X}_t^y$ . En utilisant le lemme 1 comme précédemment on a:

$$\mathbf{E}(|\varepsilon_{t_{k+1}}^N|^{2p}) \leq \mathbf{E}|\varepsilon_{t_k}^N|^{2p} (1 + C \Delta t)$$

puis,

$$\mathbf{E}(|\varepsilon_{t_k}^N|^{2p}) \leq C \mathbf{E}(|\varepsilon_0^N|^{2p}).$$

On obtient ensuite l'inégalité avec le sup sur  $t$  comme à la proposition 14.

□.

## 4.5 Convergence en loi du schéma.

L'objet d'une telle convergence est en général d'estimer des quantités du type  $\mathbf{E} g(X_T)$  pour une fonctionnelle  $g$ . Nous reviendrons sur les motivations d'un tel critère de convergence en 7.3 et nous donnerons les résultats connus sur le sujet. Précisons comment on mesure la vitesse de convergence en loi.

**Définition 3** *On dira que le schéma de discrétisation  $\bar{X}_t^N$  converge en loi sur  $\Gamma$  à l'ordre  $\alpha$  si: Pour une classe  $\Gamma$  de fonctions  $g$ , il existe une constante  $K_T(g)$  telle que:*

$$|\mathbf{E}g(X_T) - \mathbf{E}g(\bar{X}_T^N)| \leq K_T(g) \frac{1}{N^\alpha}. \quad (4.16)$$

Bien sûr si la fonction  $g$  est lipschitzienne on peut déduire des résultats de 4.2 la convergence à l'ordre  $\frac{1}{2}$  pour le schéma d'Euler. En réalité, en réduisant un peu la classe de fonctions, on peut voir que le schéma converge à l'ordre 1.

Précisons que le résultat suivant est très largement inspiré de Talay[82][86] et de Mikulevicius-Platen[86].

**Notations 1** *On notera  $C_{pol}^p$  la classe des fonctions de classe  $C^p$  qui sont, ainsi que leurs dérivées à croissance polynomiale, et comme précédemment  $C_b^p$  la classe des fonctions de classe  $C^p$  dont toutes les dérivées sont bornées.*

**Proposition 19** (Talay[82], Mikulevicius-Platen[86])

*Si  $b, \sigma$  sont  $C_b^4$ ,  $f$  est  $C_{pol}^4$  et si  $X_0$  a des moments de tous ordres, alors le schéma d'Euler est d'ordre 1 pour la classe  $C_{pol}^4$ :*

$$\exists C_T > 0, \quad |\mathbf{E}g(X_T) - \mathbf{E}g(\bar{X}_T^N)| \leq C_T \frac{1}{N}.$$

**Remarque 9** Il est possible de réduire les hypothèses sur les coefficients  $b$  et  $\sigma$ , par exemple pour inclure le cas holdérien avec coefficients bornés. L'ordre de convergence est alors plus faible, voir 7.3 ou Mikulevicius-Platen[86].

Avant de donner la preuve de cette proposition citons un résultat de Talay et Tubaro, dont nous reparlerons au chapitre 7, qui complète la proposition précédente et permet d'utiliser le schéma d'Euler pour avoir une approximation à un ordre supérieur en utilisant des interpolations de type Romberg.

**Proposition 20** (Talay-Tubaro[90])

Notons  $\xi_T = \mathbf{E}g(X_T) - \mathbf{E}g(\bar{X}_T^N)$  l'erreur d'approximation du schéma d'Euler. Si les coefficients  $b$ ,  $\sigma$  et  $g$  sont  $C_b^\infty$  alors on peut écrire pour tous  $k$ :

$$\xi_T = \frac{C_1}{N} + \dots + \frac{C_k}{N^k} + O\left(\frac{1}{N^{k+1}}\right). \quad (4.17)$$

De plus si

$$Z_T^N = 2\mathbf{E}g(\bar{X}_T^{2N}) - \mathbf{E}g(\bar{X}_T^N),$$

alors  $Z_T^N$  est une approximation en loi d'ordre 2:

$$|\mathbf{E}g(X_T) - Z_T^N| = O\left(\frac{1}{N^2}\right).$$

**Preuve de la proposition 19.** L'idée directrice est de transférer l'erreur d'approximation sur la solution de l'équation rétrograde de Kolmogorov. La preuve devient alors simple en utilisant des résultats de régularité classiques.

On peut aussi donner une démonstration probabiliste directe de cette propriété, mais l'avantage de cette approche est d'être utilisable pour étudier d'autres schémas plus complexes comme l'a fait Talay pour le deuxième schéma de Milshtein (voir 7.3).

On notera  $X_t^{t,x}$  la diffusion partant au temps  $t$  du point  $x$ .

Nous aurons besoin dans la preuve de quelques résultats classiques sur les diffusions que nous rappelons dans le lemme suivant.

**Lemme 3** (Kunita[82]) Supposons que  $b$  soit de classe  $C_b^2$ ,  $\sigma$  de classe  $C_b^4$  et  $g$  de classe  $C_{pol}^4$ .

Si  $L$  est l'opérateur de diffusion  $\sum_i b_i(t, x)\partial_i + \frac{1}{2} \sum_{i,j} (\sigma\sigma^*)_{i,j}(t, x)\partial_i\partial_j$ , alors la solution de

$$\begin{cases} \frac{\partial v}{\partial s}(s, x) + Lv(s, x) = 0 & 0 \leq s \leq T \\ v(T, x) = g(x). \end{cases}$$

est

$$v(t, x) = \mathbf{E}g(X_T^{t,x}).$$

De plus la fonction de  $t$ ,  $\mathbf{E}v(t, X_t)$ , est constante et la fonction de  $x$ ,  $v(t, x)$ , est de classe  $C_{pol}^4$ .

Donnons maintenant la preuve de la proposition 19: Notons  $\epsilon_T = \mathbf{E}g(X_T) - \mathbf{E}g(\bar{X}_T^N)$ . On a

$$\epsilon_T = \mathbf{E}v(0, X_0) - \mathbf{E}v(T, \bar{X}_T^N) = \mathbf{E}v(0, \bar{X}_0^N) - \mathbf{E}v(T, \bar{X}_T^N).$$

Il s'agit alors de montrer que

$$\mathbf{E}v(t_{k+1}, \bar{X}_{t_{k+1}}^N) - \mathbf{E}v(t_k, \bar{X}_{t_k}^N) = 0(\Delta t)^2. \quad (4.18)$$

Plaçons-nous en dimension 1 pour simplifier l'écriture. En appliquant la formule d'Itô pour  $t \in [t_k, t_{k+1}]$  il vient

$$\begin{aligned} \mathbf{E}v(t, \bar{X}_t^N) &= \mathbf{E}v(t_k, \bar{X}_{t_k}^N) + \mathbf{E} \int_{t_k}^t \partial_s v(s, \bar{X}_s^N) ds + \mathbf{E} \int_{t_k}^t b(\bar{X}_{t_k}^N) \partial_x v(s, \bar{X}_s^N) ds \\ &\quad + \frac{1}{2} \mathbf{E} \int_{t_k}^t \sigma^2(\bar{X}_{t_k}^N) \partial_{x^2}^2 v(s, \bar{X}_s^N) ds. \end{aligned}$$

En utilisant le fait que l'équation de Kolmogorov est vérifiée au point  $\bar{X}_s^N$  on obtient:

$$\begin{aligned} \mathbf{E}v(t, \bar{X}_t^N) &= \mathbf{E}v(t_k, \bar{X}_{t_k}^N) + \mathbf{E} \int_{t_k}^t (b(\bar{X}_{t_k}^N) - b(\bar{X}_s^N)) \partial_x v(s, \bar{X}_s^N) ds \\ &\quad + \frac{1}{2} \mathbf{E} \int_{t_k}^t (\sigma^2(\bar{X}_{t_k}^N) - \sigma^2(\bar{X}_s^N)) \partial_{x^2}^2 v(s, \bar{X}_s^N) ds. \end{aligned}$$

Il est facile de voir que les deux intégrales précédentes sont en  $0(\Delta t)^2$ . En effet si on note

$$f(\bar{X}_s^N) = (b(\bar{X}_{t_k}^N) - b(\bar{X}_s^N)) \partial_x v(s, \bar{X}_s^N),$$

la formule d'Itô donne

$$\mathbf{E}f(\bar{X}_s^N) = \mathbf{E} \int_{t_k}^s f'_x(\bar{X}_u^N) b(\bar{X}_{t_k}^N) du + \frac{1}{2} \mathbf{E} \int_{t_k}^s f''_{x^2}(\bar{X}_u^N) \sigma^2(\bar{X}_{t_k}^N) dW_u.$$

Sous les hypothèses du théorème il est facile de voir que  $\mathbf{E}f(\bar{X}_s^N) = 0(\Delta t)$ , ce qui montre que la première intégrale est en  $0(\Delta t)^2$ .

On procède de même pour la deuxième intégrale, ce qui donne (4.18) puis le résultat cherché.

□



## Chapitre 5

### Etude du schéma de Milshtein.





## 5.1 Introduction

Pour approcher la solution de l'E.D.S

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt, \quad t \in [0, T],$$

Milshtein[74] a proposé un nouveau schéma discret d'approximation afin d'améliorer la vitesse de convergence quadratique du schéma d'Euler. Plaçons-nous pour le moment en dimension 1. Si l'on considère une subdivision régulière  $0 = t_0 < t_1 < \dots < t_N = T$  de l'intervalle  $[0, T]$ , de pas  $\Delta t = \frac{T}{N}$ , alors le schéma de Milshtein est défini par  $\bar{X}_0^N = X_0$  et

$$\bar{X}_{t_{k+1}}^N = \bar{X}_{t_k}^N + \sigma(\bar{X}_{t_k}^N)(W_{t_{k+1}} - W_{t_k}) + b(\bar{X}_{t_k}^N)\Delta t + \frac{1}{2}\sigma(\bar{X}_{t_k}^N)\sigma'(\bar{X}_{t_k}^N)((W_{t_{k+1}} - W_{t_k})^2 - \Delta t).$$

Il converge plus rapidement que le schéma d'Euler puisque Milshtein a montré la convergence quadratique du schéma au sens suivant:

$$\exists C > 0, \forall k = 0..N, \mathbf{E}|X_{t_k} - \bar{X}_{t_k}^N|^2 \leq C (\Delta t)^2. \quad (5.1)$$

La raison en est que tous les termes d'ordre  $O(\Delta t)$  du développement de Taylor apparaissent dans ce schéma, ce qui n'était pas le cas pour le schéma d'Euler.

De plus la convergence trajectorielle de ce schéma a été obtenue par Talay[83b] via une réduction aux équations différentielles ordinaires (par le résultat de Doss[77]-Sussman[78]-Mcshane[76]):

Il existe une fonctionnelle  $F$  continue

$$F : \mathbf{R} \times C_o(\mathbf{R}_+) \rightarrow C(\mathbf{R}_+), \text{ telle que}$$

$$\forall u \in C_o(\mathbf{R}_+), F(x, u)(t) = h(D^u(t), u(t)),$$

où  $h : \mathbf{R}^2 \rightarrow \mathbf{R}$  est donnée par une équation différentielle faisant intervenir  $\sigma$ , et  $D_u$  est la solution d'une équation différentielle ordinaire faisant intervenir  $b$ ,  $\sigma$ ,  $h$  et la fonction  $u(t)$ .

La solution de l'E.D.S s'écrit alors sous la forme

$$X_t = F(X_0, W)(t), \quad \forall t \in [0, T], \text{ p.s.}$$

Talay[83b] obtient la convergence trajectorielle du schéma de Milshtein  $\bar{X}_t^u$  correspondant à la solution de

$$X_t^u = F(X_0, u)(t), \quad \forall t \in [0, T],$$

par exemple dès que  $u$  est de variation quadratique finie,  $b$  et  $\sigma$  uniformément lipschitziens, et  $\sigma$  de classe  $C^3$ . On peut même inclure les fonction de variation cubique nulle avec un peu plus de régularité. On trouve en particulier dans Talay[83b] une estimation de la vitesse de convergence p.s du schéma de Milshtein au sens suivant:

$$\forall \alpha < \frac{1}{2}, \quad N^\alpha \sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N| \xrightarrow{p.s} 0. \quad (5.2)$$

Une telle estimation provient, dans l'approche utilisée par Talay[83b], du module de continuité du brownien.

Nous allons voir que (5.2) peut être améliorée par une approche différente, i.e par une généralisation de (5.1) à la norme  $L^p$ , et qu'en fait

$$\forall \alpha < 1, \quad N^\alpha \sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N| \xrightarrow{p.s} 0. \quad (5.3)$$

L'approche utilisée par Talay permet de plus de montrer une propriété de *robustesse* qui traduit le fait heuristique suivant: Si l'on perturbe "très légèrement" le brownien, alors le schéma de Milshtein converge vers une solution "proche" de la solution exacte. Une telle propriété de robustesse pour le schéma d'Euler est discutée en 5.4 et 8.2.

## 5.2 Cas de la dimension 1.

Pour démontrer (5.3), on procède comme pour le schéma d'Euler: On va définir le schéma de Milshtein entre les temps  $t_k$  et  $t_{k+1}$  comme une diffusion, puis obtenir une estimation  $L^p$  de l'erreur par la formule d'Itô et le lemme de Gronwall, et enfin la convergence p.s par Borel-Cantelli.

Les résultats suivants sont donnés en dimension 1, nous donnerons ensuite une généralisation à la dimension  $d$ .

**Définition 4** On (re)définit le schéma de Milshtein par  $\bar{X}_0^N = X_0$  et  $\forall t \in [t_k, t_{k+1}]$ ,

$$\bar{X}_t^N = \bar{X}_{t_k}^N + \sigma(\bar{X}_{t_k}^N)(W_t - W_{t_k}) + b(\bar{X}_{t_k}^N)(t - t_k) + \sigma(\bar{X}_{t_k}^N)\sigma'(\bar{X}_{t_k}^N) \int_{t_k}^t (W_s - W_{t_k})dW_s. \quad (5.4)$$

On vérifie facilement que cette définition coïncide avec celle de Milshtein aux temps  $t_k$ .

Nous aurons besoin par la suite de quelques hypothèses de régularité sur les coefficients  $b$  et  $\sigma$ . On supposera que

- (**H<sub>1</sub>**)  $b$  et  $\sigma$  sont deux fois continûment dérivables avec des dérivées bornées.
- (**H<sub>2</sub>**)  $\exists p > 0, \mathbf{E}|X_0|^{4p} < +\infty$ .

Il est clair qu'il y a alors existence et unicité pour l'E.D.S et que la solution vérifie de plus les propriétés (Friedman[75])  $\forall q \leq 2p$ ,

$$\mathbf{E}|X_t|^{2q} \leq C_p \quad \text{et} \quad \mathbf{E}|X_t - X_s|^{2q} \leq C_p |t - s|^q.$$

On supposera par commodité que  $\Delta t \leq 1$ , ce qui n'est pas très restrictif dans la pratique.

Le résultat suivant généralise les résultats de Milshtein[74] et partiellement Talay[83b], on a la



**Proposition 21.** *Sous les hypothèses  $(H_1)$  et  $(H_2)$ ,*

$$\forall t \in [0, T], \mathbf{E}(|X_t - \bar{X}_t^N|^{2p}) \leq C(\Delta t)^{2p}, \quad (5.5)$$

$$\mathbf{E}(\sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N|^{2p}) \leq C(\Delta t)^{2p-1}. \quad (5.6)$$

*Les constantes  $C$  ne dépendent que de  $T, p, \mathbf{E}|X_0|^{4p}$  et des majorants des dérivées d'ordre 1 et 2 de  $b$  et  $\sigma$ .*

*Si de plus  $(H_2)$  est vérifiée pour tous  $p$ , alors*

$$\forall \alpha < 1, N^\alpha \sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N| \xrightarrow{p.s.} 0. \quad (5.7)$$

**Proposition 22.** *Dans le cas  $b(x) = \frac{x}{2}$  et  $\sigma(x) = x$  on a*

$$\forall \alpha > 1, \limsup_N N^\alpha |X_T - \bar{X}_T^N| = +\infty \text{ p.s.}$$

**Remarque 10** *On dira que la vitesse de convergence p.s du schéma de Milstein, d'après la formule (5.7) et la proposition 22, est en  $\frac{1}{N}$ . A titre de comparaison on a vu que la vitesse de convergence du schéma d'Euler est en  $\frac{1}{\sqrt{N}}$ .*

**Remarque 11** *Il est en fait possible d'obtenir*

$$\mathbf{E}(\sup_{t \in [0, T]} |X_t - \bar{X}_t^N|^{2p}) \leq C(\Delta t)^{2p-1/2},$$

*ce qui améliore légèrement (5.6). Toutefois la preuve étant assez fastidieuse et le résultat ne présentant pas beaucoup plus d'intérêt, nous ne donnons pas la preuve.*

**Remarque 12** *Le schéma discret  $(\bar{X}_{t_k}^N)_k$  défini par Milstein vérifie évidemment les estimations précédentes. En pratique c'est ce schéma discret qui sera utilisé. Si, comme pour le schéma d'Euler, on veut définir le schéma par interpolation linéaire entre ces points, alors la vitesse de convergence p.s est inférieure à  $\frac{1}{\sqrt{N}}$ . Il suffit pour s'en convaincre de prendre  $\sigma = 1$  et  $b = 0$ .*

Nous utiliserons dans la démonstration le résultat général suivant qui permet de passer de la convergence simple dans  $L^p$  à la convergence uniforme p.s.

**Proposition 23** *Soit un schéma de discrétisation  $\{\bar{X}_{t_k}^N\}_{k=0..N}$  tel que*

$$\forall p > 0, \exists C_p, \forall k = 0..N, \|X_{t_k} - \bar{X}_{t_k}^N\|_{L^{2p}} \leq C_p \frac{1}{N^\alpha},$$

*alors on a convergence uniforme avec probabilité 1 à l'ordre  $\alpha^-$ :*

$$\forall \epsilon > 0, N^{\alpha-\epsilon} \sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N| \xrightarrow{p.s.} 0.$$

**Preuve.** Notons

$$\xi_N = N^{\alpha-\epsilon} \sup_{k \in [0, N]} |X_{t_k} - \bar{X}_{t_k}^N|.$$

On a

$$|\xi_N|^{2p} \leq N^{2p(\alpha-\epsilon)} \sum_{k=0}^N |X_{t_k} - \bar{X}_{t_k}^N|^{2p}$$

et ainsi

$$\mathbb{E}|\xi_N|^{2p} \leq C_p \frac{1}{N^{2p\epsilon-1}}.$$

On choisit alors  $p$  tel que  $2p\epsilon - 1 > 1$  et on conclut par la version de Borel-Cantelli utilisée dans la partie sur le schéma d'Euler (lemme 2 de 4.2).

□

Nous aurons aussi besoin des deux lemmes suivants.

**Lemme 4.** Soit  $Y_t = Y_0 + \int_0^t A_s dW_s + \int_0^t B_s ds$  où  $B_s$  est un vecteur de  $\mathbb{R}^n$   $L^{2p}$ -continu,  $A_s$  une matrice de  $\mathbb{R}^{n \times d}$  à  $L^{2p}$ -continu, et  $W_t$  un brownien  $d$  dimensionnel. Alors  $Y_t$  est  $L^2$ -continu, vérifie

$$\mathbb{E}|Y_t|^{2p} \leq \mathbb{E}|Y_0|^{2p} + C \int_0^t \mathbb{E}\{|Y_s|^{2p} + |A_s|^{2p}\} + |\mathbb{E}(|Y_s|^{2p-2} Y_s \cdot B_s)| ds ,$$

et

$$\mathbb{E}|Y_t|^{2p} \leq \mathbb{E}|Y_0|^{2p} + C \int_0^t \mathbb{E}\{|Y_s|^{2p} + |A_s|^{2p} + |B_s|^{2p}\} ds .$$

**Preuve:** La preuve suit celle du lemme 1 en 4.2, avec le même argument de localisation, que nous ne réécrivons pas ici et permet de justifier l'utilisation de la formule d'Itô qui donne:

$$\begin{aligned} |Y_t|^{2p} &= |Y_0|^{2p} + \int_0^t 2p|Y_s|^{2p-2} Y_s \cdot A_s dW_s + \int_0^t 2p|Y_s|^{2p-2} Y_s \cdot B_s ds \\ &\quad + \int_0^t (p|Y_s|^{2p-2} |A_s|^2 + 2p(p-1)|Y_s|^{2p-4} |Y_s \cdot A_s|^2) ds . \end{aligned}$$

puis

$$\mathbb{E}|Y_t|^{2p} \leq \mathbb{E}|Y_0|^{2p} + C \int_0^t \mathbb{E}(|Y_s|^{2p}) + |\mathbb{E}(2p|Y_s|^{2p-2} Y_s \cdot B_s)| + \mathbb{E}(|A_s|^2 |Y_s|^{2p-2}) ds .$$

En utilisant  $x^{2p-2}y^2 \leq Cx^{2p} + Cy^{2p}$  pour  $x, y \geq 0$ , on conclut. La deuxième inégalité s'obtient aisément.

□

**Lemme 5** Si  $\mathbb{E}|X_0|^{2p} < +\infty$  alors le schéma de Mil'shtein vérifie les deux inégalités suivantes:

$$\forall t \in [0, T], \quad \mathbb{E}|\bar{X}_t^N|^{2p} \leq C ,$$

$$\forall t \in [t_k, t_{k+1}], \quad \mathbb{E}|\bar{X}_t^N - \bar{X}_{t_k}^N|^{2p} \leq C(\Delta t)^p.$$

**Preuve:** Soit  $M$  un entier quelconque. On arrête le schéma en  $\eta_M = \inf\{t > 0, |\bar{X}_t^N| > M\}$ . Par abus de notation on note encore de la même façon le schéma arrêté en  $\eta_M$ . Soit  $s \in [t_k, t_{k+1}]$ , notons  $B_s = b(\bar{X}_{t_k}^N)$   $A_s = \sigma(\bar{X}_{t_k}^N) + \sigma(\bar{X}_{t_k}^N)\sigma'(\bar{X}_{t_k}^N)(W_s - W_{t_k})$  alors on a

$$\mathbf{E}|A_s|^{2p} \leq C\mathbf{E}|\sigma(\bar{X}_{t_k}^N)|^{2p} + C\mathbf{E}|\sigma(\bar{X}_{t_k}^N)\sigma'(\bar{X}_{t_k}^N)(W_s - W_{t_k})|^{2p},$$

En utilisant  $(H_1)$  il vient

$$\mathbf{E}|A_s|^{2p} \leq C(1 + \mathbf{E}|\bar{X}_{t_k}^N|^{2p}) + C(1 + \mathbf{E}|\bar{X}_{t_k}^N|^{2p})(\Delta t)^p,$$

donc

$$\mathbf{E}|A_s|^{2p} \leq C(1 + \mathbf{E}|\bar{X}_{t_k}^N|^{2p}).$$

De même

$$\mathbf{E}|B_s|^{2p} \leq C(1 + \mathbf{E}|\bar{X}_{t_k}^N|^{2p}).$$

Comme

$$\bar{X}_t^N = \bar{X}_{t_k}^N + \int_{t_k}^t A_s dW_s + \int_{t_k}^t B_s ds,$$

le lemme précédent permet alors d'obtenir pour  $t \in [t_k, t_{k+1}]$

$$\mathbf{E}|\bar{X}_t^N|^{2p} \leq \mathbf{E}|\bar{X}_{t_k}^N|^{2p}(1 + C\Delta t) + C\Delta t + C \int_{t_k}^t \mathbf{E}|\bar{X}_s^N|^{2p} ds.$$

Si on pose  $g(s) = \mathbf{E}|\bar{X}_s^N|^{2p}$  alors le lemme de Gronwall donne pour  $t \in [t_k, t_{k+1}]$ ,

$$g(t) \leq g(t_k)(1 + C\Delta t) + C\Delta t + C \int_{t_k}^t e^{C(t-s)}(C\Delta t + g(t_k)(1 + C\Delta t))ds,$$

puis

$$g(t) \leq g(t_k)(1 + C\Delta t) + C\Delta t + Ce^{C\Delta t}(g(t_k) + C)\Delta t,$$

et donc  $g(t) \leq g(t_k)(1 + C\Delta t) + C\Delta t$ . En particulier on peut obtenir

$$g(t_{k+1}) \leq g(t_k)(1 + C\Delta t) + C\Delta t.$$

De la propriété

$$\text{Si } x_{n+1} \leq (1 + a)x_n + b \text{ alors } x_n \leq (1 + a)^n x_0 + ne^{na}b,$$

il vient tout d'abord  $g(t_k) \leq C$ , puis finalement

$$\forall k, \forall t \in [t_k, t_{k+1}], g(t) \leq C(1 + C\Delta t) \leq C.$$

On peut alors faire tendre  $M$ , introduit au début de la preuve, vers l'infini, ce qui prouve la première inégalité. Pour montrer la deuxième inégalité on écrit que

$$\bar{X}_t^N - \bar{X}_{t_k}^N = \int_{t_k}^t A_s dW_s + \int_{t_k}^t B_s ds.$$

L'inégalité de B.D.G donne alors

$$\mathbf{E}|\bar{X}_t^N - \bar{X}_{t_k}^N|^{2p} \leq (t - t_k)^{p-1} \int_{t_k}^t (\mathbf{E}|A_s|^{2p} + \mathbf{E}|B_s|^{2p}) ds.$$

L'inégalité précédente permet alors de conclure.

□

**Démonstration de la proposition 21:** Notons

$$\epsilon_t = X_t - \bar{X}_t^N, \text{ et } \xi_{t,k} = \sup_{s \in [t_k, t]} \mathbf{E}|\epsilon_s|^{2p}.$$

Pour  $t \in [t_k, t_{k+1}]$  on a

$$\epsilon_t = \epsilon_{t_k} + \int_{t_k}^t (b(X_s) - b(\bar{X}_{t_k}^N)) ds + \int_{t_k}^t (\sigma(X_s) - \sigma(\bar{X}_{t_k}^N) - \sigma(\bar{X}_{t_k}^N) \sigma'(\bar{X}_{t_k}^N)(W_s - W_{t_k})) dW_s.$$

En utilisant la formule de Taylor,

$$\sigma(\bar{X}_s^N) - \sigma(\bar{X}_{t_k}^N) = (\bar{X}_s^N - \bar{X}_{t_k}^N) \sigma'(\bar{X}_{t_k}^N) + \frac{1}{2} (\bar{X}_s^N - \bar{X}_{t_k}^N)^2 \sigma''(c_k),$$

et de plus

$$\bar{X}_s^N - \bar{X}_{t_k}^N = b(\bar{X}_{t_k}^N)(s - t_k) + \sigma(\bar{X}_{t_k}^N)(W_s - W_{t_k}) + \frac{1}{2} \sigma(\bar{X}_{t_k}^N) \sigma'(\bar{X}_{t_k}^N)((W_s - W_{t_k})^2 - (s - t_k)).$$

On a ainsi

$$\begin{aligned} \sigma(\bar{X}_s^N) - \sigma(\bar{X}_{t_k}^N) &= b \sigma'(\bar{X}_{t_k}^N)(s - t_k) + \sigma \sigma'(\bar{X}_{t_k}^N)(W_s - W_{t_k}) + \frac{1}{2} \sigma \cdot (\sigma')^2(\bar{X}_{t_k}^N)((W_s - W_{t_k})^2 - (s - t_k)) \\ &\quad + \frac{1}{2} (\bar{X}_s^N - \bar{X}_{t_k}^N)^2 \sigma''(c_k). \end{aligned}$$

On a donc

$$\epsilon_t = \epsilon_{t_k} + \int_{t_k}^t A_s dW_s + \int_{t_k}^t B_s ds,$$

avec

$$A_s = \sigma(X_s) - \sigma(\bar{X}_s^N) + b \sigma'(\bar{X}_{t_k}^N)(s - t_k) + \frac{1}{2} \sigma \cdot (\sigma')^2(\bar{X}_{t_k}^N)((W_s - W_{t_k})^2 - (s - t_k)) + \frac{1}{2} (\bar{X}_s^N - \bar{X}_{t_k}^N)^2 \sigma''(c_k),$$

et

$$B_s = b(X_s) - b(\bar{X}_{t_k}^N).$$

Pour terminer la preuve on va utiliser le lemme 4. On effectue pour cela les majorations suivantes pour  $s \in [t_k, t_{k+1}]$ :

- $\mathbf{E}|A_s|^{2p} \leq C\mathbf{E}|X_s - \bar{X}_s^N|^{2p} + C\mathbf{E}|b.\sigma'(\bar{X}_{t_k}^N)|^{2p}(\Delta t)^{2p}$   
 $+ C\mathbf{E}|\sigma.(\sigma')^2(\bar{X}_{t_k}^N)((W_s - W_{t_k})^2 - (s - t_k))|^{2p} + C\mathbf{E}|\bar{X}_s^N - \bar{X}_{t_k}^N|^{4p}.$

En utilisant le lemme 5,  $(H_1)$  et  $(H_2)$  il vient alors avec  $\xi_{t,k}$  introduit précédemment

$$\forall u \in [t_k, s], \mathbf{E}|A_u|^{2p} \leq C\xi_{s,k} + C(\Delta t)^{2p}. \quad (5.8)$$

- Il sagit maintenant de majorer  $|\mathbf{E}(\epsilon_s^{2p-1} B_s)|$  afin d'utiliser la première formule du lemme 4. Pour cela on écrit que

$$|\mathbf{E}(\epsilon_s^{2p-1} B_s)| \leq |\mathbf{E}(\epsilon_s^{2p-1} C_s)| + |\mathbf{E}(\epsilon_s^{2p-1} (b(X_{t_k}) - b(\bar{X}_{t_k}^N)))|,$$

où l'on a noté  $C_s = b(X_s) - b(X_{t_k})$ . On a alors

$$|\mathbf{E}(\epsilon_s^{2p-1} B_s)| \leq |\mathbf{E}(\epsilon_s^{2p-1} C_s)| + C\mathbf{E}|\epsilon_s|^{2p} + C\mathbf{E}|\epsilon_{t_k}|^{2p}. \quad (5.9)$$

Majorons le terme  $|\mathbf{E}(\epsilon_s^{2p-1} C_s)|$ . Pour cela on va appliquer la formule d'Itô à  $f(\epsilon_s, C_s)$  avec  $f(x, y) = x^{2p-1}y$ . En appliquant une première fois la formule de Itô on a

$$C_s = \int_{t_k}^s M_u du + \int_{t_k}^s N_u dW_u,$$

avec

$$M_u = b.b'(X_u) + \frac{1}{2}\sigma^2.b''(X_u), \quad N_u = b'.\sigma(X_u).$$

Comme en outre

$$\epsilon_s = \epsilon_{t_k} + \int_{t_k}^s A_u dW_u + \int_{t_k}^s B_u du,$$

la formule d'Itô appliquée à  $f(\epsilon_s, C_s)$  donne alors

$$\begin{aligned} \epsilon_s^{2p-1} C_s &= \epsilon_{t_k}^{2p-1} C_{t_k} + (2p-1) \int_{t_k}^s \epsilon_u^{2p-2} C_u A_u dW_u + (2p-1) \int_{t_k}^s \epsilon_u^{2p-2} C_u B_u du + \int_{t_k}^s \epsilon_u^{2p-1} M_u du \\ &\quad + \int_{t_k}^s \epsilon_u^{2p-1} N_u dW_u + (2p-1)(p-1) \int_{t_k}^s \epsilon_u^{2p-3} C_u A_u^2 du + (2p-1) \int_{t_k}^s \epsilon_u^{2p-2} A_u N_u du. \end{aligned}$$

Ainsi, en remarquant que  $C_{t_k} = 0$ ,

$$|\mathbf{E}(\epsilon_s^{2p-1} C_s)| \leq C \int_{t_k}^s \{ \mathbf{E}|\epsilon_u^{2p-2} C_u B_u| + \mathbf{E}|\epsilon_u^{2p-1} M_u| + \mathbf{E}|\epsilon_u^{2p-3} C_u A_u^2| + \mathbf{E}|\epsilon_u^{2p-2} A_u N_u| \} du.$$

Or



–  $\mathbf{E}|\epsilon_u^{2p-2}C_uB_u| \leq \mathbf{E}|\epsilon_u^{2p-2}C_u^2| + C\mathbf{E}|\epsilon_u^{2p-2}C_u\epsilon_{t_k}|$ , et  $\mathbf{E}|C_u^{2p}|^{1/p} \leq C\Delta t$ , donc

$$\mathbf{E}|\epsilon_u^{2p-2}C_uB_u| \leq C\mathbf{E}|\epsilon_u^{2p}|^{1-1/p}\Delta t + (\Delta t)^{1/2}\mathbf{E}|\epsilon_u^{2p}|^{1-1/p}\mathbf{E}|\epsilon_{t_k}^{2p}|^{1/2p}.$$

Ainsi, en remarquant que  $\Delta t \leq 1$  et en introduisant  $\xi_{t,k}$  défini au début de la démonstration on a

$$\int_{t_k}^s \mathbf{E}|\epsilon_u^{2p-2}C_uB_u| \leq C\xi_{s,k}^{1-1/p}(\Delta t)^2 + (\Delta t)^{3/2}\xi_{s,k}^{(2p-1)/2p},$$

puis en utilisant  $x^{q-1}y \leq Cx^q + Cy^q$  il vient

$$\int_{t_k}^s \mathbf{E}|\epsilon_u^{2p-2}C_uB_u| \leq C\xi_{s,k} + C(\Delta t)^{2p}.$$

–  $\mathbf{E}|\epsilon_u^{2p-1}M_u| \leq C\mathbf{E}|\epsilon_u^{2p}|^{1-1/2p}\mathbf{E}|M_u^{2p}|^{1/2p} \leq C\mathbf{E}|\epsilon_u^{2p}|^{1-1/2p}$  et donc

$$\int_{t_k}^s \mathbf{E}|\epsilon_u^{2p-1}M_u| \leq C\xi_{s,k} + C(\Delta t)^{2p}.$$

–  $\mathbf{E}|\epsilon_u^{2p-3}C_uA_u^2| \leq (\Delta t)^{1/2}\mathbf{E}|\epsilon_u^{2p}|^{1-3/2p}\mathbf{E}|A_u^{2p}|^{1/p} \leq (\Delta t)^{1/2}\xi_{s,k}^{1-3/2p}\sup_u \mathbf{E}|A_u^{2p}|^{1/p}$ .

Puis en utilisant (5.8) et  $xy^{2p-3}z^2 \leq C(x^{2p} + y^{2p} + z^{2p})$ , il vient

$$\int_{t_k}^s \mathbf{E}|\epsilon_u^{2p-3}C_uA_u^2| \leq C\xi_{s,k} + C(\Delta t)^{2p}.$$

–  $\mathbf{E}|\epsilon_u^{2p-2}A_uN_u| \leq \mathbf{E}|\epsilon_u^{2p}|^{1-1/p}\mathbf{E}|A_u^{2p}|^{1/2p}\mathbf{E}|\sigma(X_u)^{2p}|^{1/2p}$ , puis en utilisant  $xy^{2p-2}z \leq C(x^{2p} + y^{2p} + z^{2p})$ , il vient

$$\int_{t_k}^s \mathbf{E}|\epsilon_u^{2p-2}A_uN_u| \leq C\Delta t\xi_{s,k}^{1-2/2p} \sup_{u \in [t_k, s]} \mathbf{E}|A_u^{2p}|^{1/2p}$$

$$\int_{t_k}^s \mathbf{E}|\epsilon_u^{2p-2}A_uN_u| \leq C\xi_{s,k} + C(\Delta t)^{2p}.$$

On obtient alors avec les majorations précédentes

$$|\mathbf{E}\epsilon_s^{2p-1}B_s| \leq C\xi_{s,k} + C(\Delta t)^{2p}. \quad (5.10)$$

On peut maintenant appliquer le lemme 4 et on obtient avec (5.8), (5.9) et (5.10),

$$\forall t \in [t_k, t_{k+1}], \forall u \in [t_k, t], \mathbf{E}|\epsilon_u|^{2p} \leq \mathbf{E}|\epsilon_{t_k}|^{2p} + C(\Delta t)^{2p+1} + C \int_{t_k}^u \xi_{s,k} ds,$$

et ainsi

$$\forall t \in [t_k, t_{k+1}], \xi_t \leq \xi_{t_k} + C(\Delta t)^{2p+1} + C \int_{t_k}^t \xi_{s,k} ds.$$

Le lemme de Gronwall donne alors, après simplifications,

$$\xi_{t,k} \leq \xi_{t_k}(1 + C\Delta t) + C(\Delta t)^{2p+1}.$$

De la propriété

$$\text{Si } x_{n+1} \leq (1 + a)x_n + b \text{ alors } x_n \leq (1 + a)^n x_0 + ne^{na}b,$$

il vient tout d'abord

$$\exists C > 0, \forall k = 0..N, \mathbb{E}|\epsilon_{t_k}|^{2p} \leq C(\Delta t)^{2p},$$

puis

$$\exists C > 0, \forall t \in [0, T], \mathbb{E}|\epsilon_t|^{2p} \leq C(\Delta t)^{2p},$$

ce qui montre (5.5).

□

Pour montrer (5.6) et (5.7) on utilise la proposition 23 ainsi que l'argument de sa démonstration.

□

**Démonstration de la proposition 22:** On a déjà obtenu un résultat similaire pour le schéma d'Euler et la méthode est identique. Prenons  $T = 1$ , on écrit que

$$\tilde{X}_1^N = \prod_{i=1}^N (1 + \Delta W_i + \frac{1}{2}(\Delta W_i)^2) \text{ avec } \Delta W_i = W_{t_i} - W_{t_{i-1}},$$

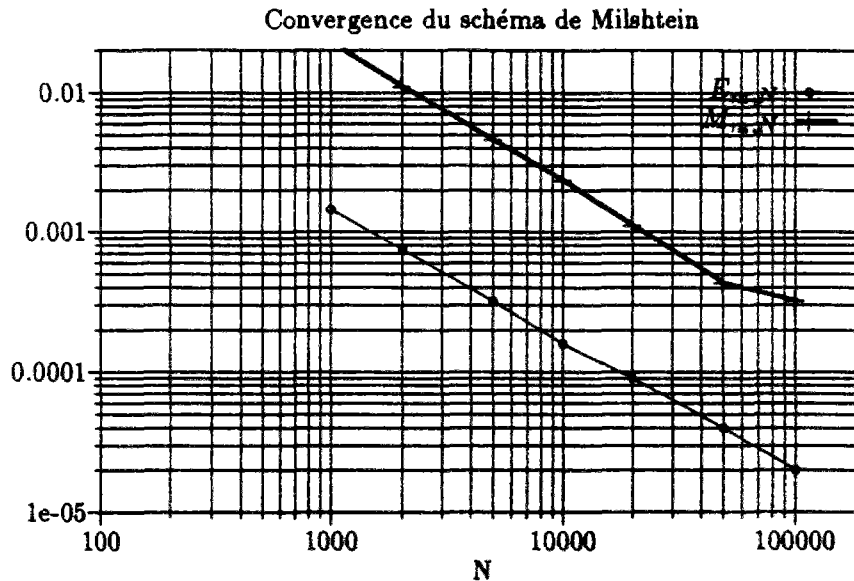
puis on obtient avec les mêmes justifications que pour le schéma d'Euler en 4.2 proposition 15,

$$\text{Log } \tilde{X}_1^N = W_1 + O(\sum_{i=1}^N (\Delta W_i)^3).$$

On conclut alors de la même façon.

□

**Un exemple numérique:** Nous reprenons l'exemple 1 de 4.3.1. Avec les mêmes conventions de notation on obtient la courbe suivante:



On constate que la convergence est bien en  $\frac{1}{N}$ .

### 5.3 Cas de la dimension n

Le cas de la dimension n, par la complication des notations, est plus difficile à écrire. De plus, et c'est le point le plus important, le résultat précédent ne se généralise bien que si la condition classique de commutativité, que nous rappelons par la suite, est vérifiée.

Précisons les notations. Considérons l'E.D.S

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt, \quad t \in [0, T],$$

où  $\sigma$  est une matrice  $\mathbf{R}^{n \times d}$ ,  $b$  un vecteur de  $\mathbf{R}^n$ , et  $(W_t) = (W_t^{(j)}, j = 1..d)$  un brownien d-dimensionnel issu de zéro. On notera

- $\sigma_j$  le  $j^{\text{ème}}$  vecteur colonne de  $\sigma$ .
- $\partial \sigma_j$  la matrice dérivée  $\partial \sigma_j = (\frac{\partial \sigma_{i,l}}{\partial x_k})_{i,k}$ .

On rappelle la définition suivante :

**Définition 5** On dira que la matrice  $\sigma$  vérifie la condition de commutativité si

$$\forall j, k = 1..d, \quad \partial \sigma_j \sigma_k = \partial \sigma_k \sigma_j,$$

qui s'écrit aussi

$$\forall i, j, k, \sum_{r=1}^n \frac{\partial \sigma_{i,j}}{\partial x_r} \sigma_{r,k} = \sum_{r=1}^n \frac{\partial \sigma_{i,k}}{\partial x_r} \sigma_{r,j}.$$

C'est sous cette condition, qui traduit le fait que l'algèbre de Lie engendrée par les  $\sigma_i$  est commutative, qu'il existe une fonctionnelle  $F$  continue telle que

$$X_t = F(X_0, W)(t) \text{ p.s.},$$

comme nous l'avons vu dans l'introduction.

Une telle condition intervient très naturellement pour discrétiser des E.D.S multidimensionnelles comme nous le verrons en partie 7.2. Schématiquement on peut dire que si cette condition de commutativité est vérifiée alors on peut écrire des schémas de discrétisation qui améliorent la vitesse de convergence quadratique du schéma d'Euler.

Dans le cas contraire, qui est le cas le plus courant, ce n'est à priori pas possible si l'on se restreint aux schémas faisant appel à des v.a facilement simulables. Voir les chapitres 7 et 8 pour plus d'informations.

Nous allons voir que l'on peut donner les mêmes résultats pour le schéma de Milshtein que dans le cas de la dimension 1, et ceci sans condition de commutativité en modifiant légèrement la définition classique. Mais cette condition sera nécessaire pour faire une simulation numérique.

Donnons la définition du schéma de Milshtein:

**Définition 6** Le schéma de Milshtein est (re)défini par  $\bar{X}_0^N = X_0$  et  $\forall t \in [t_k, t_{k+1}]$ ,

$$\bar{X}_t^N = \bar{X}_{t_k}^N + \sigma(\bar{X}_{t_k}^N)(W_t - W_{t_k}) + b(\bar{X}_{t_k}^N)(t - t_k) + \sum_{j,l=1}^d (\partial \sigma_j \sigma_l)(\bar{X}_{t_k}^N) \int_{t_k}^t (W_s^{(j)} - W_{t_k}^{(j)}) dW_s^{(l)}. \quad (5.11)$$

Pour la simulation il est nécessaire de faire disparaître l'intégrale stochastique au profit de variables aléatoires plus facilement simulables. C'est l'objet de la proposition suivante.

**Proposition 24** Si la condition de commutativité est vérifiée alors pour  $t \in [t_k, t_{k+1}]$ ,

$$\begin{aligned} \bar{X}_t^N = \bar{X}_{t_k}^N + \sigma(\bar{X}_{t_k}^N)(W_t - W_{t_k}) + (b(\bar{X}_{t_k}^N) - \frac{1}{2} \sum_{j=1}^d (\partial \sigma_j \sigma_j)(\bar{X}_{t_k}^N))(t - t_k) \\ + \frac{1}{2} \sum_{j,l=1}^d (\partial \sigma_j \sigma_l)(\bar{X}_{t_k}^N)(W_t^{(j)} - W_{t_k}^{(j)})(W_t^{(l)} - W_{t_k}^{(l)}). \end{aligned}$$

**Preuve:** Il suffit d'utiliser la formule d'intégration par partie.

□

Donnons maintenant le résultat de convergence p.s qui généralise la proposition 21:

**Proposition 25** *Sous les hypothèses  $(H_1)$  (i.e: toutes les dérivées partielles d'ordre 1 et 2 sont continues et bornées) et  $(H_2)$  on a,*

$$\forall t \in [0, T], \mathbf{E}(|X_t - \bar{X}_t^N|^{2p}) \leq C(\Delta t)^{2p}, \quad (5.12)$$

$$\mathbf{E}(\sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N|^{2p}) \leq C(\Delta t)^{2p-1}. \quad (5.13)$$

*Si de plus  $(H_2)$  est vraie pour tous  $p$ , alors*

$$\forall \alpha < 1, N^\alpha \sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N| \xrightarrow{p.s.} 0. \quad (5.14)$$

**Preuve:** La démonstration est analogue au cas de la dimension 1, mais les notations sont beaucoup plus lourdes. On utilise le lemme 4 qui a été écrit en dimension quelconque, et la suite des calculs est similaire.

□

**Remarque 13** *Bien que le résultat ci-dessus soit vrai sans condition de commutativité, le schéma de Milshtein n'est pas utilisé en pratique si cette dernière n'est pas vérifiée, car on ne sait pas simuler exactement les intégrales stochastiques dans la définition (5.11) du schéma de Milshtein. Voir les chapitres 7 et 8 à ce sujet.*

**Conclusion.** Nous avons constaté que pour améliorer la vitesse de convergence trajectorielle du schéma d'Euler, on peut utiliser le schéma de Milshtein qui converge trajectoriellement en  $1/N$ . Toutefois si l'on veut se restreindre à des schémas facilement simulables alors l'amélioration précédente n'est pertinente que si la condition de commutativité est vérifiée.

Nous verrons en partie 7.2 que cette situation particulière est en fait tout à fait générale.

**Convergence en loi du schéma de Milshtein.** Nous ne donnons pas de résultat de convergence car le schéma de Milshtein converge à la même vitesse que le schéma d'Euler, et ne présente donc pas d'amélioration. On trouvera l'énoncé de ce résultat en 7.3.

Remarquons toutefois que ce résultat est valable sans condition de commutativité, même si l'on prend comme définition du schéma de Milshtein la définition classique, à savoir celle définie par la proposition 24. Nous en comprendrons mieux la raison en partie 7.3.

## 5.4 Liens avec l'E.D.S au sens Stratonovich.

Plaçons-nous dans le cas où la condition de commutativité est vérifiée. Comme nous l'avons indiqué en 3.1.3, quand le bruit est suffisamment régulier alors le bon modèle est donné par l'E.D.S écrite au sens stratonovich et non au sens d'Itô. Plus précisément, on sait que la solution de

$$dX_t = \sigma(X_t)\xi'(t)dt + b(X_t)dt, \quad (5.15)$$

où  $\|\xi - W\|_\infty$  est petit, est proche de la solution de l'équation stratonovich

$$dX_t = \sigma(X_t)odW_t + b(X_t)dt,$$

qui s'écrit encore

$$dX_t = \sigma(X_t)dW_t + (b + \frac{1}{2}\sigma\sigma')(X_t)dt.$$

De plus si on applique le schéma de Milshtein à l'équation

$$dX_t = \sigma(X_t)dW_t + b(X_t)dt, \tag{5.16}$$

avec une trajectoire approximante du brownien régulière notée  $\xi$  alors on résoud

$$dX_t = \sigma(X_t)\xi'(t)dt + (b - \frac{1}{2}\sigma\sigma')(X_t)dt,$$

qui d'après (5.15) est proche de la solution de

$$dX_t = \sigma(X_t)odW_t + b(X_t)dt,$$

quand  $\xi$  est proche de  $W$ .

Ceci permet de mieux comprendre la propriété de robustesse du schéma de Milshtein évoquée à la fin de 5.1.

Il est clair que par ailleurs le schéma d'Euler ne possède pas la même propriété de robustesse. Le même raisonnement montre en effet que le schéma d'Euler appliqué à (5.16) pour une trajectoire approximante régulière résoud (5.15). Dans ce cas le schéma d'Euler converge vers la mauvaise solution, i.e la solution au sens stratonovich. •

Nous reviendrons sur ce point en 8.2.



## Chapitre 6

# Des schémas de discrétisation à pas adaptés.

.





## 6.1 Un résultat général de convergence.

Nous nous intéressons maintenant à une classe plus générale de schémas où le pas de discrétisation n'est plus fixe comme précédemment mais aléatoire, afin par exemple de permettre un contrôle spatial du brownien ou encore de la diffusion elle-même.

Nous allons prendre une subdivision de temps d'arrêts  $\tau_0 = 0 < \tau_1 \dots < \tau_k \dots$  recouvrant l'intervalle  $[0, T]$  et construire un schéma d'Euler défini en ces temps d'arrêts. Sous des hypothèses assez générales sur la suite de temps d'arrêts et les coefficients de la diffusion, nous allons obtenir un résultat de la forme:

$$\forall p \geq 1, \exists C_p > 0, \mathbf{E} \left( \sup_{t \in [0, T]} |X_t - \bar{X}_t^\delta|^{2p} \right) \leq C_p \delta^p, \quad (6.1)$$

où  $\delta$  représente le "pas de discrétisation spatial", homogène à un temps. Nous obtiendrons alors la convergence trajectorielle uniforme ainsi qu'une estimation de la vitesse de convergence comme nous l'avons fait au chapitre 4 pour le schéma d'Euler. Nous appliquerons alors ce résultat aux cas suivants:

- Quand la subdivision  $\tau_0 = 0 < \tau_1 \dots < \tau_n \dots$  est telle que

$$\tau_{k+1} = \tau_k + \inf \{ t > 0, |B_{t+\tau_k} - B_{\tau_k}| = \sqrt{\delta} \},$$

où  $|\cdot|$  désigne la norme du max. Ces temps d'arrêts contrôlent alors les accroissements du brownien entre deux points de la subdivision. Nous améliorons et étendons à la dimension  $d$  un résultat de Newton[90a] sur la convergence  $L^p$  et *p.s* d'un tel schéma, puis nous proposons une *méthode de simulation exacte de ce schéma* en dimension  $d$  en 6.2.

- Quand la subdivision  $\tau_0 = 0 < \tau_1 \dots < \tau_n \dots$  est telle que le schéma d'Euler  $\{\bar{X}_{t_k}^\delta\}$  vérifie

$$\forall k, |\bar{X}_{\tau_{k+1}}^\delta - \bar{X}_{\tau_k}^\delta| = \sqrt{\delta}.$$

On dispose alors d'un contrôle spatial de la diffusion. Nous complétons en 6.3 un résultat de Bichteler[81] en montrant la convergence  $L^p$  et *p.s* d'un tel schéma, puis en proposons une *méthode de simulation exacte* en dimension 1.

Nous présentons enfin en 6.4 quelques simulations numériques sur divers exemples.

### 6.1.1 Notations et hypothèses.

On considère une partition de temps d'arrêts  $\tau_0^\delta = 0 < \tau_1^\delta \dots < \tau_n^\delta \dots$ , à priori infinie, qui va recouvrir l'intervalle  $[0, T]$ . On définit le schéma d'Euler par

$$\forall t \in [\tau_k^\delta, \tau_{k+1}^\delta], \bar{X}_t^\delta = \bar{X}_{\tau_k^\delta}^\delta + \sigma(\tau_k^\delta, \bar{X}_{\tau_k^\delta}^\delta)(W_t - W_{\tau_k^\delta}) + b(\tau_k^\delta, \bar{X}_{\tau_k^\delta}^\delta)(t - \tau_k^\delta), \quad (6.2)$$

$$\bar{X}_0^\delta = X_0.$$

Nous conservons les hypothèses  $(H_1)$  à  $(H_4)$  introduites en 4.1 et supposées vraies *pour tous*  $T$  pour simplifier, et gardons les mêmes conventions de notation. Nous faisons de plus une hypothèse limitant les accroissements des pas de temps par un paramètre  $\delta$  qui représente le pas de discrétisation.

(H<sub>5</sub>) La suite  $(\tau_k^\delta)_k$  est une suite croissante de temps d'arrêts adaptée à  $\mathcal{F}_t$  telle que

$$\forall n > 0, \exists K_5 > 0, \forall k \in \mathbb{N}, \int_0^\infty \mathbf{P}(\tau_{k+1}^\delta - \tau_k^\delta > t \mid \mathcal{F}_{\tau_k^\delta})^{1/2} t^n dt \leq K_5 \delta^{n+1}. \quad (6.3)$$

Cette hypothèse implique en particulier que  $\forall p > 0, \exists C_p > 0, \forall k, \mathbf{E}((\tau_{k+1}^\delta - \tau_k^\delta)^p \mid \mathcal{F}_{\tau_k^\delta}) \leq C_p \delta^p$ .

On fait aussi une hypothèse de type grande déviation assurant que la suite de temps d'arrêts  $\tau_k^\delta$  ne recouvre pas trop lentement l'intervalle  $[0, T]$ .

(H<sub>6</sub>) Il existe une famille de v.a positives  $(Y^\delta)_\delta$  ayant des moments d'ordre 1 et 2, telles que

$$\exists K_6 > 0, \forall \delta > 0, \mathbf{E}(Y^\delta) \geq K_6 \delta,$$

$$\exists K_7 > 0, \forall \delta > 0, \mathbf{E}(Y^\delta) \leq K_7 \delta, \mathbf{E}(Y^\delta)^2 \leq K_7 \delta^2,$$

et vérifiant pour  $t \geq 0$ :

$$\forall \delta > 0, \forall n > 0, \mathbf{E}\{\exp(-t(\tau_n^\delta - \tau_{n-1}^\delta)) \mid \mathcal{F}_{\tau_{n-1}^\delta}\} \leq \mathbf{E} \exp(-t Y^\delta). \quad (6.4)$$

Enfin on introduit  $N$ , intuitivement proche du nombre moyen de pas nécessaires pour recouvrir l'intervalle  $[0, T]$ :

$$N = \left\lceil \frac{T}{K_6 \delta} \right\rceil. \quad (6.5)$$

On notera alors indifféremment  $\bar{X}^\delta$  ou  $\bar{X}^N$ .

### 6.1.2 Résultat de convergence.

Nous donnons maintenant le résultat de convergence  $L^p$  et  $p.s$  pour le schéma d'Euler à pas de temps aléatoires.

**Proposition 26.** *Sous les hypothèses  $(H_i)$  avec  $i = 1..6$ , on a*

$$(i) \sup_k \tau_k^\delta \geq T \quad p.s.$$

Si  $n^\delta = \inf\{k > 0, \tau_k^\delta > T\}$  alors

$$(ii) \exists C, C' > 0, \forall \delta > 0, \mathbf{P}(n^\delta > N + 2) \leq C \exp(-C' N).$$

De plus

$$(iii) \exists \delta_0, \forall \delta \leq \delta_0, \mathbf{E}\left(\sup_{t \in [0, T]} |X_t - \bar{X}_t^\delta|^{2p}\right) \leq C \delta^\gamma \quad \text{avec } \gamma = p \inf(2\beta, 1).$$

La constante  $\delta_0$  et la constante  $C$  de (iii) ne dépendent que des  $K_i$ ,  $T$ ,  $p$ , et  $\mathbf{E}|X_0|^{2p}$ .

Si de plus  $(H_3)$  est vérifiée pour tous  $p$  et si  $(H_4)$  est vérifiée pour  $\beta > \frac{1}{2}$  alors

$$(iv) \forall \epsilon > 0, N^{1/2-\epsilon} \sup_{t \in [0, T]} |X_t - \bar{X}_t^N| \xrightarrow{N \rightarrow +\infty} 0 \quad p.s.$$

**Remarque 14** Cette propriété généralise naturellement le résultat de convergence du chapitre 4 pour le schéma d'Euler à pas constants.  $(H_5)$  est en effet trivialement vérifiée, et pour  $(H_6)$  il suffit de prendre  $\tau_k^\delta = k\delta$  et  $Y^\delta = \delta$ .

Nous aurons besoin pour la démonstration de cette proposition des lemmes suivants:

**Lemme 6.** Si  $Y_t = Y_0 + \int_0^t A_s dW_s + \int_0^t B_s ds$ , avec  $A$  et  $B$   $L^{2p}$ -continus, alors pour tous temps d'arrêt  $\tau$  p.s fini,  $Y_t$  est  $L^{2p}$ -continu et

$$\mathbf{E}|Y_\tau|^{2p} \leq \mathbf{E}|Y_0|^{2p} + C \mathbf{E} \int_0^\tau \{|Y_s|^{2p} + |A_s|^{2p} + |B_s|^{2p}\} ds .$$

Preuve du lemme 6: On applique la formule d'Itô à  $Y_{t \wedge \tau}$  comme au lemme 1 de 4.2, puis on fait tendre  $t$  vers l'infini.

□

**Lemme 7.** (Variation sur l'inégalité de Cramer.)

Soit  $\mathcal{F}_n$  une filtration croissante et  $(X_n)_n$  une suite de v.a adaptée à cette filtration.

- Si il existe une v.a  $Y$  positive, d'espérance finie notée  $y$ , ayant des moments de tous ordres et telle que

$$\forall n \geq 0, \forall t > 0, \mathbf{E}^{\mathcal{F}_{n-1}} \exp(-tX_n) \leq \mathbf{E} \exp(-tY),$$

alors une inégalité de type grande déviation est vérifiée:

$$\forall \epsilon > 0, \exists K_\epsilon > 0, \forall n \geq 0, \mathbf{P}\left(\frac{1}{n} \sum_{i=1}^n X_i \leq y - \epsilon\right) \leq \exp(-n K_\epsilon).$$

- Si il existe une v.a  $Z$  positive, d'espérance finie  $z$ , ayant des moments exponentiels:

$$\exists \alpha > 0, \mathbf{E} \exp(\alpha Z) < +\infty,$$

et vérifiant

$$\forall n \geq 0, \forall t \in [0, \alpha], \mathbf{E}^{\mathcal{F}_{n-1}} \exp(tX_n) \leq \mathbf{E} \exp(tZ)$$

alors une inégalité de type grande déviation est aussi vérifiée:

$$\forall \epsilon > 0, \exists K_\epsilon > 0, \forall n \geq 0, \mathbf{P}\left(\frac{1}{n} \sum_{i=1}^n X_i \geq z + \epsilon\right) \leq \exp(-n K_\epsilon).$$

Preuve du lemme 7: Pour  $a$  positif fixé on peut écrire

$$\forall t > 0, \mathbf{P}\left(\frac{X_1 + \dots + X_n}{n} \leq a\right) = \mathbf{P}\left(\exp(-(X_1 + \dots + X_n)t) \geq \exp(-nta)\right),$$

et donc

$$\forall t > 0, \mathbf{P}\left(\frac{X_1 + \dots + X_n}{n} \leq a\right) \leq \exp(nta) \mathbf{E}(\exp(-t(X_1 + \dots + X_n))).$$

En conditionnant successivement par rapport à  $\mathcal{F}_n, \mathcal{F}_{n-1}, \dots$  il vient

$$\forall t > 0, \mathbf{P}\left(\frac{X_1 + \dots + X_n}{n} \leq a\right) \leq \exp\{-n \psi_a(t)\},$$

où l'on a noté

$$\begin{aligned} \forall t \geq 0, \psi_a(t) &= -at - \log \phi(t), \\ \phi(t) &= \mathbf{E}(\exp(-tY)). \end{aligned}$$

Ainsi

$$\mathbf{P}\left(\frac{X_1 + \dots + X_n}{n} \leq a\right) \leq \exp(-nK_a),$$

avec

$$K_a = \sup_{t>0} \psi_a(t).$$

Il reste donc à prouver que, pour  $a < y$ , la constante  $K_a$  est strictement positive. La fonction  $\psi_a(t)$  est de classe  $C^\infty$  sur  $\mathbf{R}^+$  et concave d'après l'inégalité de Hölder. Sa dérivée à l'origine étant  $y - a$ , elle possède donc bien un supremum strictement positif. La démonstration est similaire pour le deuxième point.

□

**Preuve de la proposition 26:** Le lemme 7 appliqué à  $X_n = (\tau_n^\delta - \tau_{n-1}^\delta)/\delta$  donne d'après  $(H_6)$

$$\forall \epsilon > 0, \forall \delta > 0, \exists C_{\epsilon,\delta} > 0, \forall n > 0, \mathbf{P}\left(\frac{\tau_n^\delta}{n\delta} \leq \frac{K_6}{1+\epsilon}\right) \leq \exp(-n C_{\epsilon,\delta}).$$

La constante  $C_{\epsilon,\delta}$  peut être choisie indépendamment de  $\delta$ . En effet, en suivant les notations de la démonstration du lemme 7, il faut montrer que si

$$\psi_{a,\delta}(t) = -at - \log(\mathbf{E}(\exp(-tY^\delta)))$$

et

$$K_{a,\delta} = \sup_{t>0} \psi_{a,\delta}(t)$$

alors

$$\inf_{\delta>0} K_{a,\delta} > 0.$$

Or il est facile de voir que la fonction  $\psi_{a,\delta}(t)$  vérifie, d'après l'hypothèse de majoration des moments d'ordre 1 et 2 de  $(H_6)$ ,

$$\forall \delta > 0, \psi'_{a,\delta}(0) \geq K_6 - a$$

et que  $\psi''_{a,\delta}(t)$  est uniformément bornée en  $\delta$  et au voisinage de  $t = 0$ . On a donc  $\inf_{\delta>0} K_{a,\delta} > 0$ . Ainsi

$$\forall \epsilon > 0, \exists C_\epsilon > 0, \forall \delta > 0, \forall n > 0, \mathbf{P}\left(\frac{\tau_n^\delta}{n\delta} \leq \frac{K_6}{1+\epsilon}\right) \leq \exp(-n C_\epsilon),$$

et donc

$$\exists C > 0, \forall \delta > 0, \forall k > N + 1, \mathbf{P}(\tau_k^\delta < T) \leq \exp(-k C). \quad (6.6)$$

Par Borel-Cantelli, on obtient (i). En particulier, pour  $k = N + 2$ , on obtient

$$\mathbf{P}(n > N + 2) \leq C e^{-C'N}, \quad (6.7)$$

ce qui montre (ii).

On s'inspire ensuite de la démonstration de 4.2. Pour cela notons  $\epsilon_t = X_t - \bar{X}_t^\delta$ .

**.Première étape:** On montre que

$$\exists C_1, C_2 > 0, \forall n > 0, \sup_{k \in [0, n]} \mathbf{E}(|\epsilon_{\tau_k}|^{2p}) \leq C_1 n \delta^{\gamma+1} e^{C_2 n \delta}.$$

Notons  $\mathbf{E}^k$  l'espérance conditionnelle par rapport à la filtration  $\mathcal{F}_{\tau_k}$ .

Remarquons que pour  $t \in [\tau_k, \tau_{k+1}]$ , on a:

$$\epsilon_t = \epsilon_{\tau_k} + \int_{\tau_k}^t (\sigma(s, X_s) - \sigma(\tau_k, \bar{X}_{\tau_k}^\delta)) dW_s + \int_{\tau_k}^t (b(s, X_s) - b(\tau_k, \bar{X}_{\tau_k}^\delta)) ds.$$

Le lemme 6 appliqué à  $Y_u = \epsilon_{\tau_k+u}$  pour  $u \in [0, t - \tau_k]$ , avec  $(H_1)$  et  $(H_4)$ , donne en suivant la démonstration de 4.2:

$$\mathbf{E}^k(|\epsilon_t|^{2p}) \leq |\epsilon_{\tau_k}|^{2p} + C \mathbf{E}^k \int_{\tau_k}^{\tau_{k+1}} \{|\epsilon_{\tau_k}|^{2p} + |X_s - X_{\tau_k}|^{2p} + |\bar{X}_s^\delta - \bar{X}_{\tau_k}^\delta|^{2p} + |\tau_{k+1} - \tau_k|^{2p\beta}\} ds.$$

De plus, en utilisant  $(H_5)$ ,

$$\mathbf{E}^k \int_{\tau_k}^{\tau_{k+1}} |\epsilon_{\tau_k}|^{2p} ds \leq |\epsilon_{\tau_k}|^{2p} \int_0^{+\infty} \mathbf{P}^k(\tau_{k+1} - \tau_k > s)^{1/2} ds \leq C \delta |\epsilon_{\tau_k}|^{2p},$$

puis

$$\mathbf{E}^k \int_{\tau_k}^{\tau_{k+1}} |X_s - X_{\tau_k}|^{2p} ds \leq \int_0^{+\infty} \mathbf{P}^k(\tau_{k+1} - \tau_k > s)^{1/2} (\mathbf{E}|X_{(s+\tau_k)} - X_{\tau_k}|^{4p})^{1/2} ds,$$

et enfin en utilisant  $(P_2)$  de la partie 4.1 et  $(H_5)$ ,

$$\mathbf{E}^k \int_{\tau_k}^{\tau_{k+1}} |X_s - X_{\tau_k}|^{2p} ds \leq C \delta^{p+1}.$$

On procède de même pour le terme  $\mathbf{E}^k \int_{\tau_k}^{\tau_{k+1}} |\bar{X}_s^\delta - \bar{X}_{\tau_k}^\delta|^{2p} ds$  car  $\bar{X}_t^\delta$  est solution de l'E.D.S (6.2).

Enfin on a par  $(H_5)$ ,

$$\mathbf{E}^k \int_{\tau_k}^{\tau_{k+1}} |\tau_{k+1} - \tau_k|^{2p\beta} ds = (2p\beta + 1) \mathbf{E}^k \int_0^\infty 1_{[0, \tau_{k+1} - \tau_k]}(t) t^{2p\beta} dt \leq C \delta^{2p\beta+1}.$$

Ainsi,

$$\mathbf{E}(|\varepsilon_{\tau_{k+1}}|^{2p}) \leq \mathbf{E}|\varepsilon_{\tau_k}|^{2p} (1 + C\delta) + C\delta^{\gamma+1}.$$

On conclut comme en 4.2.

**.Deuxième étape:** On montre qu'il existe  $\delta_0$  tel que pour  $\delta \leq \delta_0$ ,

$$\mathbf{E}(\sup_{t \in [0, T]} |\varepsilon_t|^{2p}) \leq C\delta^\gamma.$$

On écrit

$$\varepsilon_t = \int_0^t \sum_{k=0}^{\infty} (\sigma(s, X_s) - \sigma(\tau_k, \bar{X}_{\tau_k}^\delta) \mathbf{1}_{] \tau_k, \tau_{k+1} ]}(s)) dW_s + \int_0^t \sum_{k=0}^{\infty} (b(s, X_s) - b(\tau_k, \bar{X}_{\tau_k}^\delta) \mathbf{1}_{] \tau_k, \tau_{k+1} ]}(s)) ds$$

Comme en 4.2 on obtient

$$\begin{aligned} \mathbf{E}(\sup_{t \in [0, T]} |\varepsilon_t|^{2p}) &\leq C \mathbf{E} \left\{ \int_0^T \sum_{k=0}^{2N-1} |\sigma(s, X_s) - \sigma(\tau_k, \bar{X}_{\tau_k}^\delta)|^{2p} \mathbf{1}_{] \tau_k, \tau_{k+1} ]}(s) ds \right\} \\ &\quad + C \mathbf{E} \left\{ \int_0^T \sum_{k=0}^{2N-1} |b(s, X_s) - b(\tau_k, \bar{X}_{\tau_k}^\delta)|^{2p} \mathbf{1}_{] \tau_k, \tau_{k+1} ]}(s) ds \right\} \\ &\quad + C \mathbf{E} \left\{ \int_0^T \sum_{k=2N}^{\infty} |\sigma(s, X_s) - \sigma(\tau_k, \bar{X}_{\tau_k}^\delta)|^{2p} \mathbf{1}_{] \tau_k, \tau_{k+1} ]}(s) ds \right\} \\ &\quad + C \mathbf{E} \left\{ \int_0^T \sum_{k=2N}^{\infty} |b(s, X_s) - b(\tau_k, \bar{X}_{\tau_k}^\delta)|^{2p} \mathbf{1}_{] \tau_k, \tau_{k+1} ]}(s) ds \right\}. \end{aligned}$$

Les deux premières intégrales se majorent facilement par  $C\delta^\gamma$  en utilisant l'étape 1 comme en 4.2. Pour les deux autres on remarque que

$$\begin{aligned} &\mathbf{E} \left\{ \int_0^T \sum_{k=2N}^{\infty} (|\sigma(s, X_s) - \sigma(\tau_k, \bar{X}_{\tau_k}^\delta)|^{2p} + |b(s, X_s) - b(\tau_k, \bar{X}_{\tau_k}^\delta)|^{2p}) \mathbf{1}_{] \tau_k, \tau_{k+1} ]}(s) ds \right\} \\ &\leq \mathbf{E} \left\{ \int_0^T \sum_{k=2N}^{\infty} (|\sigma(s, X_s) - \sigma(\tau_k, \bar{X}_{\tau_k}^\delta)|^{2p} + |b(s, X_s) - b(\tau_k, \bar{X}_{\tau_k}^\delta)|^{2p}) \mathbf{1}_{] \tau_k, \tau_{k+1} ]}(s) \mathbf{1}_{\{\tau_k < T\}}(s) ds \right\}, \\ &\leq C \int_0^T \sum_{k=2N}^{\infty} (\mathbf{E}(|\sigma(s, X_s) - \sigma(\tau_k, \bar{X}_{\tau_k}^\delta)|^{4p} + |b(s, X_s) - b(\tau_k, \bar{X}_{\tau_k}^\delta)|^{4p}) \mathbf{1}_{] \tau_k, \tau_{k+1} ]}(s))^{1/2} \mathbf{P}(\tau_k < T)^{1/2} ds. \end{aligned}$$

Notons  $I$  cette intégrale. On a, en procédant comme en 4.2, et d'après l'étape 1 et (6.6)

$$I \leq C \int_0^T \sum_{k=2N}^{\infty} (\delta^p + k\delta^{\gamma+1} e^{C_2 k \delta}) e^{-K_8 k} ds.$$

On choisit alors  $\delta_0$  tel que

$$\exists C_4 > 0, \forall \delta \leq \delta_0, \quad C_2 \delta - K_8 < -C_4 < 0.$$

Il vient alors

$$I \leq C\delta^\gamma.$$

Ce qui démontre (iii).

**.Troisième étape:** On conclut comme pour le schéma d'Euler en 4.2. pour obtenir (iv).

□

## 6.2 Un schéma avec contrôle spatial du brownien

Newton[90a] utilise une subdivision de temps d'arrêts contrôlant les accroissements du brownien en dimension 1. Nous procédons de même et, dans le cas du schéma d'Euler, améliorons son résultat et l'étendons à la dimension  $d$ . Comme précédemment, nous obtenons des estimations de convergence  $L^p$  et  $p.s.$  Enfin nous étudions la simulation numérique de ce schéma et présentons un exemple numérique.

### 6.2.1 Résultat de convergence.

On considère la subdivision  $\tau_0 < \tau_1 \dots < \tau_n \dots$  définie par:

$$\tau_0 = 0 \text{ et } \tau_{k+1} = \tau_k + \min_{i \in [1..d]} \{ \inf \{ t > 0, |B_{t+\tau_k}^i - B_{\tau_k}^i| = \sqrt{\delta} \} \}.$$

On a omis le  $\delta$  dans  $\tau_n^\delta$  pour alléger l'écriture. On introduit aussi  $N$  qui correspond intuitivement le nombre moyen de pas de subdivision:

$$N = \left\lceil \frac{T}{C_d \delta} \right\rceil,$$

où  $C_d$  représente  $\mathbf{E}\tau_1$  quand  $\delta = 1$  (cf. partie 2.3.).

Les temps d'arrêts  $\tau_k$  sont tels que, conditionnellement au passé avant  $\tau_k$ ,  $\tau_{k+1} - \tau_k$  représente le temps de sortie pour le brownien d'un cube de côté  $2\sqrt{\delta}$  centré en  $B_{\tau_k}$ . Une telle subdivision est donc plus riche en informations que la subdivision usuelle de pas constant car elle permet un contrôle des accroissements du brownien entre les points de la subdivision.

**Définition 7** *Le schéma d'Euler E.C.B (Euler avec Contrôle spatial du Brownien) est défini par  $\bar{X}_0^N = X_0$  et:*

$$\forall t \in [\tau_k, \tau_{k+1} \wedge T], \bar{X}_t^N = \bar{X}_{\tau_k}^N + \sigma(\tau_k, \bar{X}_{\tau_k}^N)(W_t - W_{\tau_k}) + b(\tau_k, \bar{X}_{\tau_k}^N)(t - \tau_k) \quad (6.8)$$

Par convention le schéma est arrêté au temps  $T$ . Comme au chapitre 4, on définit un schéma d'Euler analogue par interpolation linéaire, noté  $\tilde{X}_t^N$ , entre les points de la subdivision. On a la

**Proposition 27.** *Sous les hypothèses  $(H_1)$  à  $(H_4)$  de 4.2 on a:*

$$(i) \quad \forall k \geq 0, \tau_k < +\infty,$$



$$(ii) \lim_{k \rightarrow \infty} \tau_k \geq T \text{ p.s.}$$

Le temps de recouvrement  $n = \inf\{k, \tau_k \geq T\}$  vérifie

$$(iii) N \leq \mathbf{E}n < N + 2,$$

ainsi qu'une inégalité de type grande déviation:

$$(iv) \exists C, C' > 0, \mathbf{P}(\{n < N\} \cup \{n > N + 2\}) \leq C \exp(-C'N).$$

De plus

$$(v) \mathbf{E}(\sup_{t \in [0, T]} |X_t - \bar{X}_t^N|^{2p}) \leq C \delta^\gamma \text{ avec } \gamma = p(2\beta \wedge 1).$$

La constante  $C$  ne dépendant que de  $K_1, K_2, K_4, T, p$  et  $\mathbf{E}|X_0|^{4p}$ .

Enfin si de plus  $\gamma > 1$  et  $(H_3)$  est vérifiée pour tous  $p$ , alors il y a convergence p.s à l'ordre  $\frac{1}{2} - \epsilon$ :

$$(vi) \forall \epsilon > 0, N^{1/2-\epsilon} \sup_{t \in [0, T]} |X_t - \bar{X}_t^N| \xrightarrow{p.s} 0.$$

**Remarque 15** Ce résultat améliore celui de Newton[90a] valable en dimension 1:

$$\limsup_{\delta \rightarrow 0} \sup_{k \leq N} \mathbf{E}\left\{\left(\frac{|X_{t_k} - \bar{X}_{t_k}^N|}{\sqrt{\delta}}\right)^p\right\} < +\infty.$$

**Démonstration de la proposition 27:** Par la proposition 26, il suffit que la suite de temps d'arrêts  $(\tau_k)_k$  vérifie les hypothèses  $(H_5)$  et  $(H_6)$ . La suite  $\tau_{k+1} - \tau_k$  est une suite iid d'espérance proportionnelle à  $\delta$ ,

$$\mathbf{E}(\tau_{k+1} - \tau_k) = C_d \delta,$$

et ayant des moments de tous ordres (cf partie 2.3).  $(H_5)$  est vérifiée d'après la proposition 9 de 2.4.3. De plus  $(H_6)$  est vérifiée en prenant  $Y^\delta = \tau_1$ . La proposition 26 permet alors d'obtenir (i), (ii), (v) et (vi).

On remarque que le théorème d'arrêt appliqué à la martingale  $\tau_k - k C_d \delta$  donne après localisation

$$\mathbf{E} \tau_n = C_d \delta \mathbf{E}n.$$

De  $\tau_n \geq T$  on obtient  $\mathbf{E}n \geq N$ . Puis en écrivant  $\tau_n = \tau_{n-1} + (\tau_n - \tau_{n-1})$  il vient de la même façon  $\mathbf{E}n < N + 2$ , ce qui prouve (iii). Enfin pour (iv) on remarque que la première partie provient de (6.7). Puis,  $\tau_1$  admettant des moments exponentiels positifs d'après la proposition 8 de 2.4.1 (qui était donnée en dimension 1 et est à fortiori vraie en dimension plus grande), et en utilisant le lemme 7 pour  $Z = \tau_1$  et  $X_n = (\tau_n - \tau_{n-1})$  il vient

$$\forall \epsilon > 0, \exists C_\epsilon > 0, \forall n > 0, \mathbf{P}\left(\frac{\tau_n}{n\delta} \geq \frac{K_6}{1-\epsilon}\right) \leq e^{-C_\epsilon n}.$$

En remarquant que la constante  $C_\epsilon$  peut être choisie indépendamment de  $\delta$  par scaling, on a

$$\mathbf{P}(n < N) \leq C e^{-C'n},$$

ce qui prouve (iv).  $\square$

**Remarque 16** Remarquons que l'on peut inclure sans difficulté le cas plus général où

$$\tau_{k+1} = \tau_k + \min_{i \in [0..d]} \inf \{ t > 0, |B_t^i - B_{\tau_k}^i| = f_k \sqrt{\delta} \} \}$$

$f_k$  étant  $\mathcal{F}_{\tau_k}$  adaptée et vérifiant  $\exists C_1, C_2, \forall k > 0, 0 < C_1 < f_k < C_2$ . Il suffit en effet d'introduire les deux suites

$$\tau_k^{(1)} = \min_{i \in [0..d]} \inf \{ t > 0, |B_{\tau_k+t}^i - B_{\tau_k}^i| = C_1 \sqrt{\delta} \},$$

$$\tau_k^{(2)} = \min_{i \in [0..d]} \inf \{ t > 0, |B_{\tau_k+t}^i - B_{\tau_k}^i| = C_2 \sqrt{\delta} \}.$$

Il est facile de reprendre le raisonnement précédent en "coinçant"  $\tau_{k+1} - \tau_k$  entre  $\tau_k^{(1)}$  et  $\tau_k^{(2)}$ .

Contrairement au schéma d'Euler classique, on obtient les mêmes estimations de vitesse de convergence pour le schéma d'Euler interpolé linéairement défini en 4.2. On a en effet la

**Proposition 28** Le schéma d'Euler interpolé linéairement  $\tilde{X}_t^N$ , défini en 4.2, vérifie, sous les mêmes hypothèses, les résultats (i) à (iv) de la proposition 27.

**Démonstration:** Seule la propriété (v) reste à prouver. Notons  $\varepsilon_t = \tilde{X}_t^N - \bar{X}_t^N$  la différence entre les deux schémas d'Euler. On a:

$$\forall t \in [\tau_k, \tau_{k+1}[ , \varepsilon_t = \sigma(\tau_k, \bar{X}_{t_k}^N) (W_t - W_{\tau_k} - \frac{t - t_k}{t_{k+1} - t_k} (W_{t_{k+1}} - W_{t_k}))$$

et donc

$$\forall t \in [\tau_k, \tau_{k+1}[ , |\varepsilon_t| \leq |\sigma(\tau_k, \bar{X}_{t_k}^N)| (|W_t - W_{\tau_k}| + |W_{t_{k+1}} - W_t|).$$

Ainsi

$$\sup_{t \in [\tau_k, \tau_{k+1}]} |\varepsilon_t| \leq 2\sqrt{\delta} |\sigma(\tau_k, \bar{X}_{t_k}^N)|.$$

Puis,  $\sigma(\tau_k, \bar{X}_{t_k}^N)$  étant uniformément borné en norme  $L^p$ , on obtient le résultat.

$\square$

**Remarque 17** Un résultat de convergence analogue peut être obtenu pour le schéma de Milshtein défini le long de cette subdivision aléatoire, qui généralise les résultats du chapitre 5, et permet d'obtenir une convergence en  $1/N$  quand la condition de commutativité est vérifiée.

### 6.2.2 Simulation numérique: Un exemple d'application.

Pour simuler le schéma E.C.B il faut savoir simuler le couple  $(\tau_{k+1} - \tau_k, B_{\tau_{k+1}} - B_{\tau_k})$  à chaque étape. Mais c'est précisément le problème qui a été résolu dans le chapitre sur les temps de sortie du mouvement brownien; nous avons vu un algorithme assez rapide permettant de faire une telle simulation en partie 2.3.

Quelle utilisation peut-on faire de ce schéma ? Il nous semble qu'un tel schéma permet de mieux rendre compte, en temps très petit, du comportement aléatoire de la diffusion. On trouve par exemple dans Newton[90a] un résultat d'optimalité de vitesse de convergence d'un schéma analogue, comportant plus de termes que ce schéma d'Euler, et basé sur une telle subdivision aléatoire. On peut aussi penser utiliser ce schéma quand une simulation spatiale du brownien, dans l'esprit du chapitre 2, est en même temps souhaitable.

Remarquons toutefois que l'on ne peut *observer* la convergence trajectorielle d'une tel schéma vers la solution directement; il faudrait en effet effectuer un raffinement de la trajectoire brownienne. Or nous ne savons faire un tel raffinement, i.e simuler un pont d'un brownien conditionné à rester à l'intérieur d'une cube.

Nous avons cependant testé le schéma E.C.B sur l'exemple déjà utilisé en 4.3 pour le schéma d'Euler classique, afin de faire quelques comparaisons numériques. Nous conservons les mêmes notations; si

$$e_\delta = \sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^\delta|$$

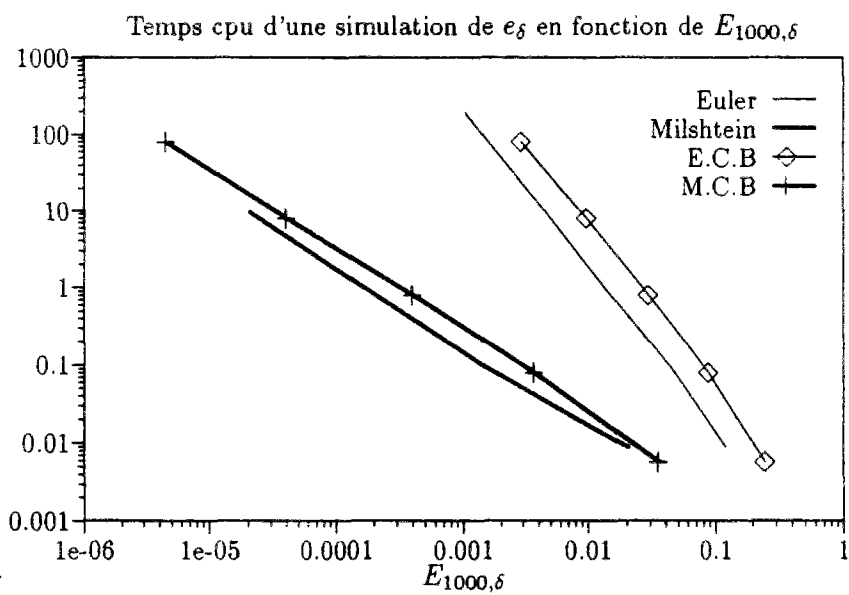
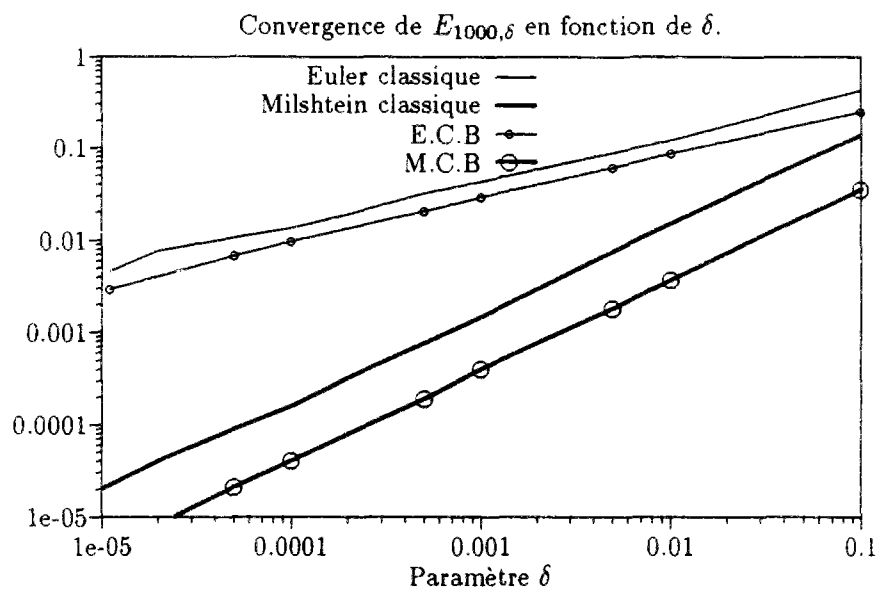
et si  $(e_\delta^1, \dots, e_\delta^m)$  sont  $m$  réalisations indépendantes de l'erreur d'approximation  $e_\delta$ , on estime

$$E_{m,\delta} = \frac{e_\delta^1 + \dots + e_\delta^m}{m}$$

La première courbe de la page suivante montre que la convergence du schéma E.C.B est très similaire, sur cet exemple, à celle du schéma d'Euler classique.

Il en est de même pour le schéma de type Milshtein M.C.D, basé sur la même subdivision de temps d'arrêts, comme indiqué en remarque 16.

La deuxième courbe compare le temps cpu moyen d'une simulation de  $e^\delta$  en fonction d'une précision donnée, ici l'erreur  $E_{m,\delta}$ . Ce temps cpu est proportionnel à la dimension de la simulation, i.e au nombre de nombres pseudo-aléatoires utilisés pour faire une simulation de  $e^\delta$ . Là encore, les performances sont similaires à celles des schémas à pas fixes.



### 6.3 Un schéma avec contrôle spatial de la diffusion.

L'idée d'un tel schéma est d'utiliser une subdivision aléatoire qui permette une approximation spatiale de la diffusion. Plus précisément on va construire une subdivision  $\tau_0^\delta = 0 < \tau_1^\delta < \dots < \tau_n^\delta$  et un schéma d'approximation  $\bar{X}_t^\delta$  tels que

$$\forall k, |\bar{X}_{\tau_{k+1}^\delta}^\delta - \bar{X}_{\tau_k^\delta}^\delta| = \sqrt{\delta},$$

où  $\delta$  représente le pas de discrétisation spatial homogène à un temps.

Ce schéma a été introduit par Bichteler[81], dans un cadre plus général, et il en a obtenu la convergence *p.s* quand la suite  $\delta_n$  de pas de temps est diadique.

Un tel schéma est à priori plus difficile à simuler qu'un schéma d'Euler classique, et c'est sans doute pour cela qu'il n'a pas été plus étudié. Nous allons montrer que:

- Ce schéma converge dans  $L^p$  et *p.s* vers la diffusion comme le schéma d'Euler-Maruyama sous quelques hypothèses supplémentaires.
- Il est possible, en dimension 1, de simuler un tel schéma bien qu'il ne fasse pas appel à des v.a élémentaires comme des gaussiennes.

Un tel schéma peut permettre d'observer des propriétés fines des diffusions; nous pensons qu'un tel schéma possède les mêmes propriétés de dégénérescence, de récurrence, d'explosion... que la diffusion correspondante, et peut donc en permettre une étude fine.

Tous les résultats suivants de convergence sont donnés en dimension  $d$  mais nous verrons que nous ne savons faire la simulation d'un tel schéma qu'en dimension 1 essentiellement.

#### 6.3.1 Résultat de convergence.

**Définition 8** *Le schéma Euler E.C.D (Euler avec Contrôle spatial de la Diffusion) est défini par*

$$\bar{X}_0^\delta = X_0, \tau_0 = 0,$$

$$\forall t \in [\tau_k, \tau_{k+1} \wedge T], \bar{X}_t^\delta = \bar{X}_{\tau_k}^\delta + b(\bar{X}_{\tau_k}^\delta)(t - \tau_k) + \sigma(\bar{X}_{\tau_k}^\delta)(W_t - W_{\tau_k}),$$

et

$$\tau_{k+1} = \tau_k + \inf\{t > 0, |b(\bar{X}_{\tau_k}^\delta)t + \sigma(\bar{X}_{\tau_k}^\delta)(W_{t+\tau_k} - W_{\tau_k})| = \sqrt{\delta}\}.$$

Ainsi défini ce schéma vérifie pour  $\tau_{k+1} \leq T$ ,  $|\bar{X}_{\tau_{k+1}^\delta}^\delta - \bar{X}_{\tau_k^\delta}^\delta| = \sqrt{\delta}$  et  $\delta$  apparait comme un pas de discrétisation spatial homogène à un temps. Remarquons que nous avons omis le symbole  $\delta$  dans  $\tau_k^\delta$  pour alléger l'écriture.

Comme pour le schéma d'Euler classique nous faisons des hypothèses assurant l'existence et l'unicité de la solution de l'E.D.S considérée, à savoir les hypothèses  $(H_1)$ ,  $(H_3)$ ,  $(H_4)$  de 4.1 et nous supposons aussi que les coefficients sont bornés:

$$(H_2') \quad \exists M_b, M_\sigma > 0, \forall t \in [0, T], \forall x \in \mathbf{R}^n, |b(t, x)| \leq M_b, |\sigma(t, x)| \leq M_\sigma.$$

On fait de plus une hypothèse supplémentaire permettant de contrôler les accroissements des temps aléatoires de discrétisation, à savoir

$$(A). \quad \exists \sigma_0 > 0, \forall x \in \mathbf{R}^d, |\sigma(x)| \geq \sigma_0,$$

ou bien

$$(B). \quad \exists b_0 > 0, \forall x \in \mathbf{R}^d, |b(x)| \geq b_0.$$

Nous reviendrons sur le caractère un peu restrictif de ces hypothèses.

Posons  $N = T/\delta$  qui représente approximativement le nombre de pas de discrétisation nécessaire pour recouvrir l'intervalle  $[0, T]$ . On peut alors énoncer le résultat de convergence:

**Proposition 29.** *Sous les hypothèses  $(H_1)$ ,  $(H_3)$ ,  $(H_4)$  de 4.1,  $(H'_2)$  et  $(A)$  on a*

$$(i) \quad \forall k > 0, \tau_k < +\infty \text{ p.s.}$$

$$(ii) \quad \lim_{k \rightarrow \infty} \tau_k \geq T \text{ p.s.}$$

Si  $n^\delta = \inf\{t > 0, \tau_k^\delta > T\}$  alors on a

$$(iii) \quad \exists C, C', C'' > 0, \forall \delta > 0, \mathbf{P}(n^\delta > CN) \leq C' \exp(-C'' N).$$

De plus

$$(iv) \quad \forall T > 0, \mathbf{E}(\sup_{t \in [0, T]} |X_t - \bar{X}_t^\delta|^{2p}) \leq C\delta^p,$$

La constante  $C$  ne dépendant que de  $K_1, K_2, K_4, T, p, \mathbf{E}|X_0|^{4p}$  et  $\sigma_0$ .

Si de plus  $(H_3)$  est vérifiée pour tous  $p$ , alors il y a convergence p.s.:

$$(v) \quad \forall \epsilon > 0, N^{1/2-\epsilon} \sup_{t \in [0, T]} |X_t - \bar{X}_t^N| \xrightarrow{p.s.} 0.$$

Enfin, si on suppose  $(B)$  au lieu de  $(A)$ , les mêmes propriétés sont vérifiées en remplaçant  $\delta$  par  $\delta^{1/2}$  dans  $(iv)$  et  $N$  par  $N^{1/2}$  dans  $(v)$ .

**Remarque 18** *Il doit être possible d'affaiblir l'hypothèse de bornitude sur  $b$  et  $\sigma$  pour inclure le cas linéaire, bien que nous n'ayons pas trouvé de preuve dans ce cas. Il doit aussi être possible d'affaiblir l'hypothèse  $(A)$  qui semble un peu restrictive.*

**Preuve de la proposition 29.** Il suffit de vérifier  $(H_5)$  et  $(H_6)$  d'après la proposition 26.  $(H_5)$  est obtenue par la proposition 9 de 2.4.3. Pour vérifier  $(H_6)$  on introduit

$$\nu_k^\delta = \inf\{t > 0, |M_\sigma(W_{\tau_k+t} - W_{\tau_k}) + M_b t| = \sqrt{\delta}\}.$$

Plaçons-nous pour le moment en dimension 1. On a

$$\mathbf{E}\{ \exp(-t(\tau_{k+1} - \tau_k)) | \mathcal{F}_{\tau_k} \} \leq \mathbf{E} \exp(-t\nu_1^\delta). \quad (6.9)$$

En effet  $\nu_k^\delta$  a même loi que  $\nu_1^\delta$ . De plus si  $b(\bar{X}_{\tau_k}^\delta)$  et  $\sigma(\bar{X}_{\tau_k}^\delta)$  sont de même signes  $\tau_{k+1} - \tau_k \geq \nu_k^\delta$  et le résultat est évident. S'ils sont de signes opposés, alors le raisonnement est similaire en utilisant cette fois  $\eta_k^\delta = \inf\{t > 0, |M_\sigma(W_{\tau_k+t} - W_{\tau_k}) - M_b t| = \sqrt{\delta}\}$ , qui a même loi que  $\nu_k^\delta$ . Ce qui prouve (6.9).

De plus

$$\begin{aligned} \exists C > 0, \forall \delta > 0, \quad \mathbf{E}(\nu_i^\delta) &\leq C\delta, \quad \mathbf{E}(\nu_i^\delta)^2 \leq C\delta^2, \\ \exists C > 0, \quad \mathbf{E}(\nu_1^\delta) &\geq C\delta. \end{aligned} \quad (6.10)$$

Les deux premières inégalités proviennent de la proposition 9 de 2.4.3. Pour montrer (6.10) on remarque que, d'après (2.4) proposition 8, on a

$$\begin{aligned} \mathbf{E}(\tau_1^\delta) &= \frac{\sqrt{\delta}}{M_b} th\left(\frac{M_b \sqrt{\delta}}{M_\sigma^2}\right) \mathbf{1}_{\{M_b \sqrt{\delta} M_\sigma^{-2} > 1\}} + \mathbf{1}_{\{M_b \sqrt{\delta} M_\sigma^{-2} \leq 1\}} \\ &\geq \frac{\sqrt{\delta}}{M_b} (th(1) + C \frac{M_b \sqrt{\delta}}{M_\sigma^2}) \\ &\geq C\delta. \end{aligned}$$

( $H_6$ ) est donc vérifiée, ce qui achève la preuve. Pour la dimension  $d > 1$  la preuve est similaire en raisonnant composante par composante.

Enfin, si on suppose l'hypothèse (B) au lieu de (A), on utilise la proposition 9 de 2.4.3 et le reste de la preuve est similaire.

□

### 6.3.2 Simulation numérique: Un exemple d'application.

Comme pour le schéma d'Euler E.C.B le problème de la simulation a été résolu dans le chapitre sur la simulation contrôlée du mouvement brownien avec drift en 2.4.2 car, à chaque étape, les coefficients du brownien avec drift à simuler sont constants.

Nous ne savons malheureusement pas simuler ce schéma en général en dimension strictement plus grande que 1, comme nous l'avons vu en 2.4.3.

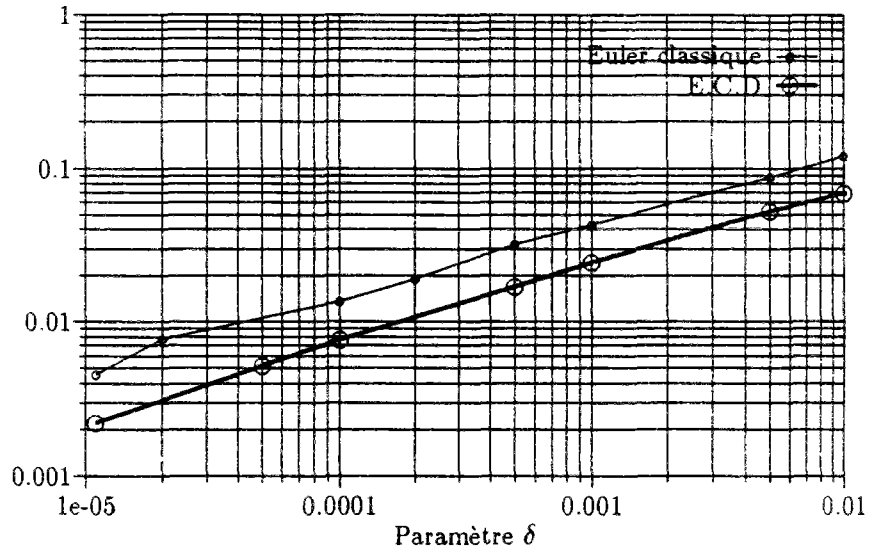
Donnons un exemple simple de vérification de la convergence de ce schéma. Pour cela nous reprenons le même exemple qu'en 4.3 et 6.2 et, avec les mêmes notations qu'en 6.2, nous obtenons les deux courbes de la page suivante.

La première courbe montre que la convergence du schéma E.C.D est similaire à celle du schéma d'Euler classique.

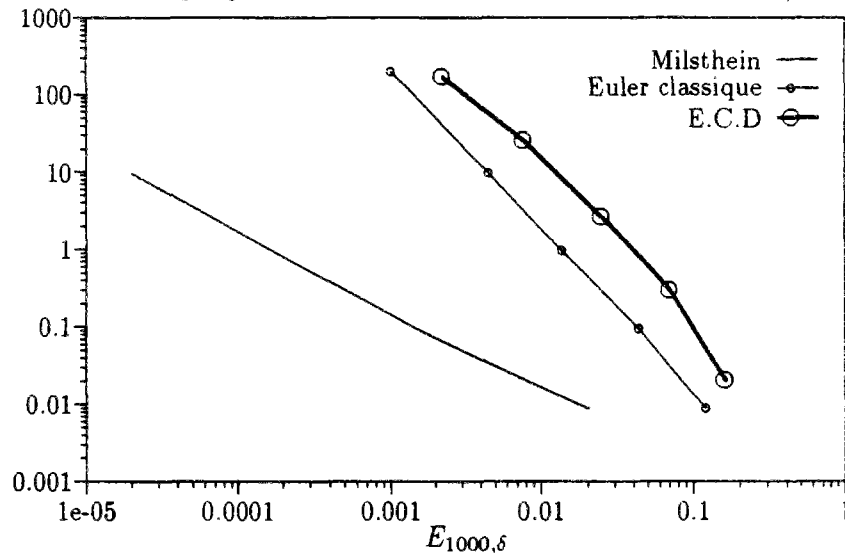
La deuxième courbe compare le temps cpu de simulation de ces deux schémas en fonction d'une précision donnée, ici  $E_{m,\delta}$ . Comme rappelé en 6.2, ce temps cpu est proportionnel à la dimension de la simulation. Nous avons placé aussi la courbe correspondante pour le schéma de Milshtein classique à titre de comparaison.

Comme pour le schéma E.C.B, les performances sont très similaires.

Convergence de  $E_{1000,\delta}$  en fonction de  $\delta$ .



Temps cpu d'une simulation de  $e_\delta$  en fonction de  $E_{1000,\delta}$





## 6.4 Quelques simulations numériques

Supposons que les coefficients  $b$  et  $\sigma$  soient localement lipschitziens en  $x$ , i.e:

$$\forall N > 0, \exists K_N > 0, \forall x, y \text{ t.q. } |x| \leq N, |y| \leq N, |\sigma(x) - \sigma(y)| + |b(x) - b(y)| \leq K_N |x - y|. \quad (6.11)$$

Il n'y a à priori plus croissance linéaire en  $x$  et on peut montrer qu'il existe une solution forte pour l'E.D.S, définie de façon unique sur  $[0, e(\omega)[$ , où  $e$  est appelé le temps d'explosion de la diffusion. Si on introduit

$$\eta_M = \inf\{t > 0, |X_t| \geq M\},$$

alors le temps d'explosion  $e$  peut être vu comme la limite p.s de  $\eta_M$  quand  $M$  tend vers l'infini (Voir par exemple Ikeda-Watanabe[81]).

Un exemple simple est l'E.D.S en dimension 1:

$$dX_t = X_t(1 + X_t^2)dt + (1 + X_t^2)dW_t.$$

Les coefficients ne sont pas globalement lipschitziens et la solution

$$X_t = \tan(W_t),$$

est définie jusqu'au temps d'explosion aléatoire

$$e(\omega) = \inf\{t > 0, |W_t| = \frac{\pi}{2}\}.$$

Enfin précisons que l'on connaît des critères permettant de savoir à priori si il y a ou non explosion de la solution, voir par exemple Arnold[74] ou Rogers-Williams[88].

On peut remarquer que tous les schémas de discrétisation à pas constant ont un *temps d'explosion infini*, et sont donc mal indiqués pour approcher une diffusion de temps d'explosion fini.

Ce n'est pas le cas pour le schéma E.C.D. Considérons en effet l'exemple suivant:

$$dX_t = X_t^2 dt, \quad X_0 = \frac{1}{c}$$

de solution

$$X_t = \frac{1}{c - t}, \quad t \in [0, c[$$

définie jusqu'au temps d'explosion déterministe  $e = c$ . Le schéma E.C.D vérifie

$$\bar{X}_{\tau_k}^\delta = k\sqrt{\delta} + \frac{1}{c},$$

$$\tau_k = \sum_{i=0}^{k-1} \frac{\sqrt{\delta}}{(i\sqrt{\delta} + 1/c)^2}.$$

On peut définir pour le schéma E.C.D de façon analogue:

$$\bar{\eta}_M^\delta = \inf\{t > 0, |\bar{X}_t^\delta| \geq M\},$$

et  $\bar{e}^\delta$  comme limite p.s de  $\bar{\eta}_M^\delta$  quand  $M$  tend vers l'infini.

Le temps d'explosion  $\bar{e}^\delta$  donné par

$$\bar{e}^\delta = \sum_{i=0}^{\infty} \frac{\sqrt{\delta}}{(i\sqrt{\delta} + 1/c)^2},$$

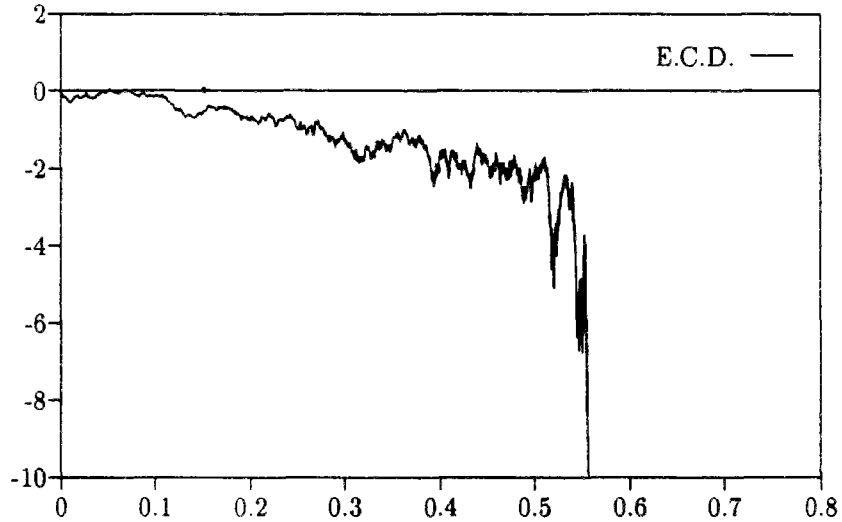
vérifie, en encadrant la série précédente par un intégrale,

$$0 \leq \bar{e}^\delta - e \leq e^2 \sqrt{\delta}. \quad (6.12)$$

Ainsi, sur cet exemple déterministe, nous avons approché le temps d'explosion de la solution et mesuré la vitesse de convergence de cette approximation.

Nous n'avons pas trouvé de résultat général de convergence dans un cadre stochastique. Il doit être possible de montrer la convergence du schéma E.C.D vers la solution sur les compacts de  $[0, e \wedge \bar{e}^\delta[$  ainsi que la convergence de  $\bar{e}^\delta$  vers  $e$ .

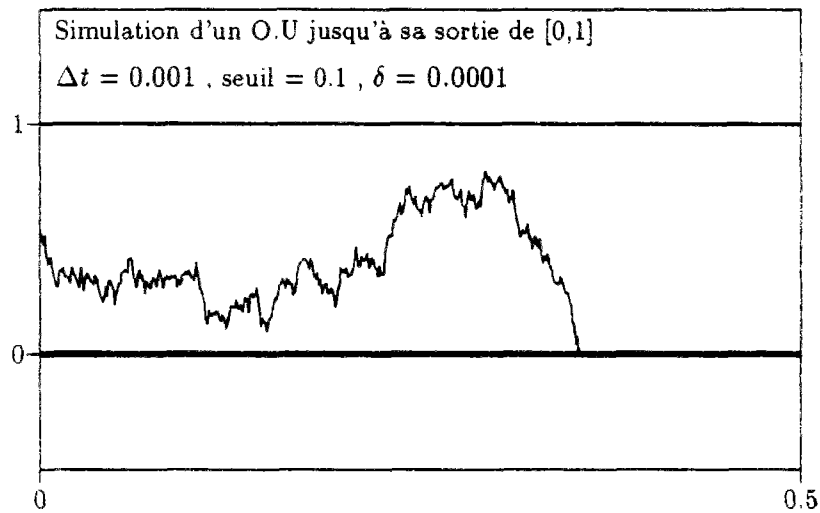
Nous donnons ici, à titre d'illustration, un exemple de simulation numérique pris parmi d'autres simulations analogues. Nous reprenons l'exemple précédent et simulons une trajectoire du schéma E.C.D. Le schéma E.C.D est arrêté quand il atteint en valeur absolue une constante  $M$  fixée. Sur la figure suivante apparaît la trajectoire du schéma E.C.D pour  $M = 20$  et  $\delta = (0.05)^2$ . La solution exacte aux points de discrétisation est aussi représentée, et est ici confondue avec le schéma.



Nous ne savons pas si le schéma E.C.D permet d'étudier les diffusions qui explosent dans le cas général.

Dans l'exemple suivant nous approchons une diffusion jusqu'à sa sortie d'un domaine, dans l'esprit de 2.5.2. Considérons une diffusion, ici un processus d'O.U partant de 0.5, dont nous voulons simuler une "bonne" approximation p.s jusqu'à sa sortie de l'intervalle  $[0, 1]$ . Pour accélérer la convergence nous utilisons tout d'abord le schéma de Milstein avec un pas  $\Delta t$  puis, quand la distance à la frontière notée  $dist$ , est inférieure à un *seuil* fixé par avance, nous utilisons le schéma E.C.D avec un pas inférieur ou égale à  $\delta \wedge dist$ . Enfin la procédure est arrêtée quand le schéma atteint la frontière.

Cette méthode est illustrée par la figure suivante:



On peut ainsi envisager des méthodes mixtes utilisant des schémas à pas fixes, puis à pas aléatoires de type E.C.D quand une simulation plus fine est nécessaire.

## **Chapitre 7**

### **Panorama des travaux existants.**

Nous présentons dans cette partie un panorama des travaux existants sur la discrétisation des équations différentielles stochastiques. Cette présentation poursuit deux objectifs:

- Donner une vision d'ensemble des travaux sur le sujet par un exposé synthétique contenant de nombreuses références bibliographiques.
- Permettre au non spécialiste une utilisation immédiate pour les applications sans besoin de se référer à de nombreux articles.

Cette présentation est bien sûr subjective; nous avons choisit de présenter des résultats qui nous semblent pertinents pour les applications numériques. Il y a donc de nombreuses omissions de travaux, certes importants, mais qui nous paraissent difficilement utilisables en pratique.

On pourra se référer pour de plus amples informations aux survey de Talay[90a], Kloden-Platen[89], Clark[88], Pardoux-Talay[85]. Un ouvrage de Kloden-Platen[91] devrait aussi sortir prochainement.

Enfin précisons que cette partie peut être lue, dans une large mesure, indépendamment des chapitres précédents.

## 7.1 Introduction

Soit l'équation différentielle stochastique (E.D.S) de type Itô:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad t \in [0, T], \quad (7.1)$$

avec  $b: \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^n$ ,  $\sigma: \mathbf{R} \times \mathbf{R}^n \rightarrow \mathbf{R}^{n \times d}$  et  $(W_t)$  un brownien standard de  $\mathbf{R}^d$ . En général on considérera une subdivision régulière de l'intervalle de temps  $[0, T]$ ,  $t_0 = 0 < t_1 < \dots < t_N = T$  de pas constant  $t_{k+1} - t_k = \Delta t = \frac{T}{N}$ .

Un schéma de discrétisation sera noté  $\{\bar{X}_t^N\}$  et sera souvent défini de façon itérative par une formule du type:

$$\bar{X}_{t_{k+1}}^N = f(t_k, \bar{X}_{t_k}^N, \Delta t, W_{t_{k+1}} - W_{t_k}), \quad (7.2)$$

Un tel schéma sera appelé un schéma à pas fixe.

Le choix d'un schéma de discrétisation dépend très largement du type de problème étudié. Il y a deux grandes classes de schémas de discrétisation qui correspondent à des motivations différentes:

- Si l'on désire obtenir, pour une trajectoire du brownien donnée, la trajectoire correspondante de la diffusion, alors on utilisera un *critère d'approximation forte* (quadratique, p.s). Les travaux concernant la convergence forte (ou trajectorielle) des schémas de discrétisation sont présentés en partie 7.2.
- Si l'on désire obtenir des informations sur la loi de la diffusion, on peut alors se contenter d'utiliser un *critère d'approximation faible*. Les travaux concernant la convergence faible (ou convergence en loi) des schémas de discrétisation sont présentés en partie 7.3.

Ces deux critères d'approximation conduisent à des schémas de discrétisation distincts et à des problématiques très différentes comme nous allons le voir.

## 7.2 Approximations fortes.

Le premier critère de convergence forte utilisé a été un critère de convergence quadratique. C'est celui qui a été le plus abondamment étudié, sans doute parce qu'il est le plus facile à manipuler, et nous commençons par présenter des résultats suivant ce critère de convergence. Mais ce n'est sans doute pas le plus pertinent pour les applications. Il est à notre avis préférable pour mesurer une convergence forte d'utiliser le critère de convergence trajectorielle présenté en 7.2.2.

### 7.2.1 Convergence quadratique.

**Définition 9** On dira que le schéma de discrétisation  $\{\bar{X}_{t_k}^N\}$  converge à l'ordre  $\alpha$  en moyenne quadratique si

$$\exists C_T > 0, \forall k = 0..N, (\mathbf{E}|X_{t_k} - \bar{X}_{t_k}^N|^2)^{\frac{1}{2}} \leq C_T \frac{1}{N^\alpha}. \quad (7.3)$$

Nous avons vu au chapitre 4 que Maruyama[55] a obtenu la convergence à l'ordre  $\frac{1}{2}$  du schéma Euler, qui est l'analogue du schéma d'Euler pour les équations différentielles ordinaires. Pour améliorer cette vitesse de convergence, Milshtein[74] a introduit un nouveau schéma, qui porte maintenant son nom. Ce schéma converge à l'ordre 1, au moins sous la condition de commutativité sur  $\sigma$ :

$$\forall i, j, k, \sum_{r=1}^d \frac{\partial \sigma_{i,j}}{\partial x_r} \sigma_{r,k} = \sum_{r=1}^d \frac{\partial \sigma_{i,k}}{\partial x_r} \sigma_{r,j}. \quad (7.4)$$

Nous avons aussi vu que ces résultats peuvent être améliorés et généralisés à la norme  $L^p$  aux chapitres 4 et 5, puisque nous avons montré que:

- Pour le schéma d'Euler

$$\forall p \geq 1, \exists C_{p,T} > 0, \mathbf{E} \left( \sup_{t_k \in [0,T]} |X_{t_k} - \bar{X}_{t_k}^N|^{2p} \right)^{\frac{1}{2p}} \leq C_{p,T} \frac{1}{N^{\frac{1}{2}}}. \quad (7.5)$$

- Pour le schéma de Milshtein

$$\forall p \geq 1, \exists C_{p,T} > 0, \sup_{t_k \in [0,T]} (\mathbf{E}|X_{t_k} - \bar{X}_{t_k}^N|^{2p})^{\frac{1}{2p}} \leq C_{p,T} \frac{1}{N}, \quad (7.6)$$

et

$$\forall p \geq 1, \exists C_{p,T} > 0, \mathbf{E} \left( \sup_{t_k \in [0,T]} |X_{t_k} - \bar{X}_{t_k}^N|^{2p} \right)^{\frac{1}{2p}} \leq C_{p,T} \frac{1}{N^{1-\frac{1}{2p}}}. \quad (7.7)$$

Nous avons aussi constaté au chapitre 5 que si la condition de commutativité (7.4) sur  $\sigma$  n'est pas vérifiée, le schéma de Milshtein possède la même vitesse de convergence que le schéma d'Euler. On peut modifier la définition classique du schéma de Milshtein afin de retrouver les résultats (7.6) et (7.7), mais ce nouveau schéma n'est pas utilisable en pratique car on ne sait pas le simuler correctement. Voir 8.1 à ce sujet.

Une question naturelle est alors de se demander s'il est possible d'écrire des schémas de discrétisation qui améliorent d'avantage ces vitesses de convergence comme on le fait pour les équations différentielles ordinaires. Mais malheureusement il y a deux réponses possibles:

- La première réponse est non: Il n'est pas possible d'obtenir une convergence à un ordre supérieur à  $\frac{1}{2}$  dans le cas général et à 1 dans le cas où la condition de commutativité (7.4) est vérifiée... si on se limite à une classe de schémas de discrétisation *facilement simulables*. C'est l'objet du résultat de Clark et Cameron que nous rappelons par la suite.
- La deuxième réponse est oui: Il est possible d'écrire des schémas qui convergent à un ordre arbitraire ... mais interviennent alors dans la définition de ces schémas des intégrales stochastiques, dont la simulation numérique est difficile. De tels schémas de discrétisation sont alors difficilement utilisables en pratique. C'est l'objet du résultat de Platen et Wagner que nous donnons par la suite.

Afin de formaliser ces deux réponses, introduisons quelques notations.

Considérons une subdivision  $\sigma_N = (t_0 = 0 < t_1 \dots < t_N = T)$  de l'intervalle  $[0, T]$ , de pas (par exemple) régulier  $t_{k+1} - t_k = \Delta t = T/N$ . On note  $P_{\sigma_N}$  la  $\sigma$ -algèbre engendrée par les  $(W_k \frac{T}{N}, k = 0..N)$ .

**Définition 10** On dira qu'un schéma de discrétisation est  $P_{\sigma_N}$ -simulable s'il existe un algorithme fini permettant d'obtenir ce schéma à partir des seules v.a  $(W_k \frac{T}{N}, k = 0..N)$ .

Un schéma  $P_{\sigma_N}$ -simulable est donc en particulier  $P_{\sigma_N}$ -mesurable, mais le contraire n'est pas vrai.

Il faut remarquer que la plupart des schémas connus sont  $P_{\sigma_N}$ -simulables. Pourquoi se restreindre à cette classe de schémas de discrétisation?

C'est une classe raisonnablement large de schémas de discrétisation qui présentent l'avantage d'être facilement simulables puisqu'ils ne font appels qu'à des v.a du type  $W_{t_{k+1}} - W_{t_k}$ , qui sont des gaussiennes. De plus on ne dispose parfois dans certaines applications comme en filtrage, que des valeurs du brownien à des instants discrétisés, ce sont alors les seuls schémas utilisables. Malheureusement cette classe de schémas de discrétisation "facilement simulables" possède certaines limites comme l'illustre le résultat suivant de Clark et Cameron.

Remarquons que pour le critère de convergence quadratique la meilleure approximation  $L^2$  de la diffusion appartenant à la classe de schémas  $P_{\sigma_N}$ -mesurable est  $\mathbf{E}(X_T | P_{\sigma_N})$ . Donnons maintenant les deux résultats limitatifs:

**Proposition 30** (Clark-Cameron[80]).

Soient  $W_t^{(1)}$  et  $W_t^{(2)}$  deux browniens indépendants. Considérons la diffusion  $(X_t, Y_t)$  définie par

$$dX_t = Y_t dW_t^{(1)},$$

$$dY_t = dW_t^{(2)},$$

$$X_0 = Y_0 = 0.$$

Cette diffusion ne vérifie pas la condition de commutativité (7.4) sur  $\sigma$  et on a

$$\|X_T - \mathbf{E}(X_T|P_{\sigma_N})\|_{L^2} = \frac{T}{2\sqrt{N}}. \quad (7.8)$$

Le schéma d'Euler  $\{\bar{X}_t^N\}$  est presque optimal car il vérifie

$$\|X_T - \bar{X}_T^N\|_{L^2} = \frac{T}{\sqrt{2N}}. \quad (7.9)$$

**Proposition 31** (Clark-Cameron[80]).

Si  $b$  est de classe  $C_b^3$  alors la diffusion solution de

$$dX_t = b(X_t)dt + dW_t, \quad X_0 = \alpha,$$

vérifie la condition de commutativité (7.4) et on a

$$\|X_T - \mathbf{E}(X_T|P_N)\|_{L^2} = \frac{1}{N} \left\{ \frac{T^3}{12} \int_0^T \mathbf{E}(\exp(2 \int_t^T b'(X_s) ds) b'(X_s)^2) dt \right\}^{1/2} + o\left(\frac{1}{N}\right). \quad (7.10)$$

Ces deux résultats limitent donc à 1 et à  $\frac{1}{2}$  les ordres de convergence quadratique, suivant que la condition de commutativité (7.4) est respectivement vérifiée ou non, pour la classe de schémas  $P_{\sigma_N}$ -simulables.

Il est toutefois possible que, sur certains exemples, la convergence soit plus rapide; le schéma d'Euler converge à un ordre infini dans le cas  $b = 0$  et  $\sigma = 1$  !

Ces vitesses de convergence peuvent néanmoins être améliorées à condition d'utiliser d'autres schémas de discrétisation. Nous donnons en ce sens le résultat de Platen et Wagner annoncé précédemment. L'idée est de faire un développement de Taylor stochastique pour écrire des schémas à un ordre arbitraire de convergence. Pour cela introduisons les opérateurs

$$L^0 = \sum_{i=1}^n b^i \partial_i + \frac{1}{2} \sum_{i,j=1}^n (\sigma \cdot \sigma^*)_{i,j} \partial_{i,j}^2,$$

et

$$L^j = \sum_{i=1}^n \sigma_{i,j} \partial_i.$$

On notera  $dW_s^0 = ds$  et  $\sigma_{k,0} = b_k$ . De plus on introduit les intégrales stochastiques

$$\begin{aligned} I_j &= \int_{t_k}^{t_{k+1}} dW_s^{(j)} \\ I_{j_1, j_2} &= \int_{t_k}^{t_{k+1}} \int_{t_k}^{s_2} dW_{s_1}^{(j_1)} dW_{s_2}^{(j_2)}, \\ I_{j_1, j_2, j_3} &= \int_{t_k}^{t_{k+1}} \int_{t_k}^{s_3} \int_{t_k}^{s_2} dW_{s_1}^{(j_1)} dW_{s_2}^{(j_2)} dW_{s_3}^{(j_3)}. \end{aligned}$$



**Proposition 32** (Platen-Wagner[78], Platen[81])

Le schéma de discrétisation défini par  $\bar{X}_t = (\bar{X}_t^{(1)}, \dots, \bar{X}_t^{(n)})$  et

$$\bar{X}_{t_{k+1}}^{(i)} = \bar{X}_{t_k}^{(i)} + \sum_{j=1}^d \sigma_{i,j}(\bar{X}_{t_k}) I_j + \sum_{j_1, j_2=1}^d (L^{j_1} \sigma_{i,j_2})(\bar{X}_{t_k}) I_{j_1, j_2} + \sum_{j_1, j_2, j_3=1}^d (L^{j_1} L^{j_2} \sigma_{i,j_3})(\bar{X}_{t_k}) I_{j_1, j_2, j_3}$$

est un schéma d'ordre 1.5 si les coefficients sont assez réguliers.

On trouve aussi dans la même référence un schéma d'approximation à un ordre arbitraire; il suffit de pousser le développement de Taylor plus loin.

Il faut remarquer que ces schémas font appel à des v.a qui sont des intégrales stochastiques multiples que l'on ne sait pas en général simuler numériquement. On pourrait essayer de les approcher mais il est fort probable que l'on perde alors, par de nécessaires approximations, tout l'intérêt de tels schémas. Voir la partie 8.1 à ce sujet.

Il ressort de ces propositions que l'on ne peut pas améliorer d'avantage l'ordre de convergence des schémas de discrétisation, si l'on se restreint à des schémas facilement simulables, i.e  $P_{\sigma_N}$ -simulables. Le schéma de Milshtein ayant alors l'ordre de convergence "optimal" quand la condition (7.4) est vérifiée, sinon il faut se contenter du schéma d'Euler.

Néanmoins on peut espérer trouver des schémas ayant le même ordre de convergence que ces schémas, mais dont la constante  $C_T$  dans (7.3) serait plus faible. C'est une direction de recherche qui a été étudiée, on peut citer notamment les travaux de Clark[80][82][84], Newton[86][90b].

La technique utilisée consiste à obtenir des schémas asymptotiquement meilleurs, i.e minimisant la constante  $C_T$  parmi la classe des schémas  $\sigma_n$ -mesurables, par des théorèmes de type centrale limite. Ces résultats sont surtout intéressants si la filtration  $\sigma_N$  est imposée et que l'on veut simuler la "meilleure" approximation possible. Mais dans le cas général, ces résultats étant seulement asymptotiques, il n'est pas dit que d'autres schémas convergent plus rapidement pour des ordres de grandeur de  $N$  utilisés en pratique. De plus ces schémas comportent en général beaucoup plus de termes que le schéma d'Euler ou de Milshtein et la complexité de tels schémas peut parfois compenser le gain de vitesse de convergence.

Plaçons nous maintenant dans le cas où la condition de commutativité (7.4) est vérifiée. Ce sera le cas par exemple quand le brownien est monodimensionnel. Nous avons vu qu'il est alors possible d'utiliser le schéma de Milshtein pour obtenir une convergence d'ordre 1. Mais ce schéma présente l'inconvénient de faire apparaître les dérivées de  $\sigma$ , ce qui peut se révéler lourd en pratique en grandes dimensions ou quand la dérivée est particulièrement difficile à calculer.

On peut contourner cette difficulté en écrivant des schémas de type Runge-Kutta. Cela a été fait par Rumelin[82], Platen[86], Newton[91]. Nous donnons ici un résultat de Newton qui répond à la question dans le cas d'un brownien monodimensionnel.

**Définition 11** (Newton[91])

Le schéma FRKI (First order Runge-Kutta method involving Itô coefficient) est défini par  $\bar{X}_0 = X_0^N$  et :

$$\bar{X}_{t_{k+1}}^N = \bar{X}_{t_k}^N + b(\bar{X}_{t_k}^N) \Delta t - \sigma_0 \sqrt{\Delta t} + \sigma_1 (\Delta W_{k+1} + \sqrt{\Delta t}), \quad (7.11)$$

avec

$$\begin{aligned}\sigma_0 &= \sigma(\bar{X}_{t_k}^N) \\ \sigma_1 &= \sigma(\bar{X}_{t_k}^N + \frac{1}{2}\sigma_0(\Delta W_{k+1} - \sqrt{\Delta t})) \\ \Delta W_{k+1} &= W_{t_{k+1}} - W_{t_k}.\end{aligned}$$

On a alors la

**Proposition 33** (Newton[91])

Dans le cas d'un brownien monodimensionnel, si les coefficient  $b$  et  $\sigma$  sont de classe  $C_b^1$ , alors

$$\forall p > 0, \sup_{k=0..N} (N \mathbf{E}|X_{t_k} - \bar{X}_{t_k}^N|)^p < +\infty. \quad (7.12)$$

Un tel schéma possède donc le même ordre de convergence optimal que le schéma de Milshtein mais ne fait pas intervenir de dérivée. Remarquons qu'il est sans doute possible d'étendre ces résultats au cas plus général où le brownien est multidimensionnel et où la condition de commutativité (7.4) est vérifiée.

Citons aussi une autre direction de recherche qui consiste à écrire des schémas de discrétisation pour des E.D.S avec sauts. On peut citer dans ce sens Gikhmann-Skorokhod[79], Platen[82], Mikulevicius-Platen[88] et Leplingle-Ribémont[89] où est traité le cas de sauts poissoniens. Mais il n'y a pas eu à notre connaissance de simulations numériques dans ce cadre montrant la pertinence de ces méthodes.

Enfin dans le cas d'une E.D.S du second ordre, on sait que l'on peut se ramener à une E.D.S classique en incluant les dérivés dans les composantes du vecteur. On peut bien sûr écrire un schéma de discrétisation sur cette équation mais il peut être préférable d'écrire directement un schéma de discrétisation sur l'équation du deuxième ordre. Voir à ce sujet Leplingle-Ribémont[89].

**Conclusion:** Il ressort de ce panorama les constatations suivantes. Suivant le critère de convergence quadratique, deux cas distincts se présentent:

- Si la condition de commutativité (7.4) sur  $\sigma$  n'est pas vérifiée, alors on utilisera en général le schéma d'Euler, qui est un schéma d'ordre  $\frac{1}{2}$ . On ne connaît pas dans ce cas de schéma facilement simulable ayant un ordre de convergence quadratique supérieur et le schéma d'Euler a pour lui son extrême simplicité d'écriture. Voir aussi 8.1 à ce sujet.
- Si la condition de commutativité (7.4) sur  $\sigma$  est vérifiée, on peut bien sûr utiliser le schéma d'Euler mais il sera en général préférable d'utiliser le schéma de Milshtein qui est d'ordre 1. On ne sait pas véritablement écrire des schémas plus performants et facilement simulables.

Toutefois, dans le cas d'un brownien monodimensionnel, on pourra utiliser le schéma FRKI de type Runge-Kutta de Newton défini précédemment; c'est aussi un schéma d'ordre 1 mais il ne nécessite pas l'utilisation des dérivées de  $\sigma$ .

### 7.2.2 Convergence p.s ou trajectorielle.

L'idée est de donner pour une trajectoire du brownien, une bonne approximation de la trajectoire correspondante de la diffusion. Bien que cette idée semble naturellement adaptée à certaines applications comme en filtrage, ou en statistique pour observer l'influence de certains paramètres, ou encore pour l'étude expérimentale fine des trajectoire des diffusions, il existe curieusement très peu de travaux sur le sujet.

Précisons ce que l'on entend par convergence trajectorielle.

**Définition 12** *On dira que le schéma de discrétisation  $\{\bar{X}_t^N\}$  converge trajectoriellement à l'ordre  $\alpha$  si:*

$$N^\alpha \sup_{t_k \in [0, T]} |X_{t_k} - \bar{X}_{t_k}^N| \xrightarrow{N \rightarrow \infty} 0 \text{ p.s.} \quad (7.13)$$

*On dira que le schéma de discrétisation  $\{\bar{X}_t^N\}$  converge ponctuellement à l'ordre  $\alpha$  si:*

$$N^\alpha |X_T - \bar{X}_T^N| \xrightarrow{N \rightarrow \infty} 0. \quad (7.14)$$

Nous avons vu au chapitre 4 que le schéma d'Euler converge trajectoriellement à l'ordre  $\frac{1}{2} - \epsilon$  pour tous  $\epsilon > 0$  et que cette vitesse de convergence ne peut être d'ordre supérieur à  $\frac{1}{2}$ .

De plus, suivant ce critère de convergence, on retrouve la dichotomie créée par la condition de commutativité (7.4) sur  $\sigma$  vue en 7.2.1. Plus précisément:

- *Si la condition de commutativité (7.4) sur  $\sigma$  n'est pas vérifiée, alors le schéma d'Euler converge trajectoriellement vers la diffusion. On peut sans doute écrire d'autres schémas de discrétisation qui convergent trajectoriellement, mais aucun schéma  $P_{\sigma_N}$ -simulable ne permet à priori d'obtenir une convergence plus rapide que le  $\frac{1}{\sqrt{N}}$  du schéma d'Euler. Voir 8.1 sur la difficulté intrinsèque d'un simulation numérique dans ce cas.*
- *Si la condition de commutativité (7.4) sur  $\sigma$  est vérifiée, alors il est possible d'améliorer cette vitesse de convergence p.s. Nous avons vu en effet au chapitre 5 que le schéma de Milshtein converge trajectoriellement à l'ordre  $1 - \epsilon$  pour tous  $\epsilon > 0$  et que cette vitesse de convergence ne peut être d'ordre supérieur à 1 en général.*

Il est fort probable que, comme pour la convergence quadratique, on ne puisse pas faire mieux en utilisant des schémas facilement simulables, i.e  $P_{\sigma_N}$ -simulables.

Précisons que dans le cas où le brownien est monodimensionnel, Newton[91] a obtenu la convergence ponctuelle du schéma FRKI de type Runge-Kutta, introduit dans la partie précédente, et ceci au même ordre de convergence que le schéma de Milshtein. En fait il est facile de voir que la convergence est même trajectorielle par la proposition 23 du chapitre 5.2.

On trouve aussi dans Newton[91] un résultat de convergence en loi de  $N^{\frac{1}{2}} |X_T - \bar{X}_T^N|$  et  $N |X_T - \bar{X}_T^N|$  respectivement pour le schéma d'Euler et pour le schéma FRKI.

On peut définir un critère de convergence plus sévère que la convergence trajectorielle qui permet de mesurer la sensibilité du schéma de discrétisation à la précision de la simulation du brownien. Pour cela on donne la définition suivante :

**Définition 13** (Talay[83b]) *On dira que la convergence trajectorielle du schéma de discrétisation  $\{\bar{X}_T\}$  est robuste pour la classe de fonctions  $\mathbf{C}$  si:*

*Pour une certaine classe  $\mathbf{C}$  de fonctions  $u$ , contenant une version régulière du brownien standard et muni d'une certaine métrique  $d$ , le schéma de discrétisation  $\{\bar{X}_T^u\}$  obtenu en remplaçant dans la définition du schéma le brownien par la fonction  $u$ , converge vers la diffusion quand  $u$  tend vers le brownien suivant la métrique  $d$ .*

Le résultat de Talay, rappelé en 5.1, montre que pour le schéma de Milshtein on peut prendre pour  $\mathbf{C}$  la classe des fonctions de  $C_0$  de variation cubique nulle. Le schéma d'Euler n'est à priori robuste que pour une classe très restreinte, voir 8.2. à ce sujet.

Une conséquence numérique de ces propriétés peut être la "stabilité" du schéma de Milshtein. Si par simulation du brownien on commet une "petite erreur", alors la trajectoire simulée reste "proche" de la véritable solution.

**Conclusion.** Pour le critère de convergence trajectorielle on retrouve la même dichotomie que pour la convergence quadratique:

- Si la condition de commutativité (7.4) sur  $\sigma$  n'est pas vérifiée, alors le schéma d'Euler est le seul schéma  $P_{\sigma_N}$ -simulable (i.e facilement simulable) dont on connaisse la convergence trajectorielle avec une vitesse d'environ  $\frac{1}{\sqrt{N}}$ . Voir 8.1 pour compléments.
- Si la condition de commutativité est vérifiée, alors il est préférable d'utiliser par exemple le schéma de Milshtein dont la vitesse de convergence trajectorielle est en environ  $\frac{1}{N}$  et qui est robuste pour une large classe de fonctions.

Si le brownien est monodimensionnel on peut utiliser le schéma FRKI de Newton rappelé précédemment qui converge à la même vitesse que le schéma de Milshtein, mais ne fait pas intervenir les dérivées de  $\sigma$ .

Nous n'avons pas parlé des schémas à pas variables ou aléatoires car nous connaissons très peu de résultats montrant l'intérêt numérique de ces schémas en pratique. Voir toutefois 8.3 où l'utilisation de pas variable permet de contrôler l'erreur finale. Voir aussi le chapitre 6 où nous avons introduit deux schémas à pas aléatoires, permettant de contrôler les accroissements du brownien ou de la diffusion elle-même. Nous ne savons pas encore si de tels schémas peuvent être pertinent en pratique.

### 7.2.3 Le cas d'une E.D.S au sens Stratonovich.

Nous avons rappelé en 3.1 la transformation permettant de passer d'une E.D.S au sens d'Itô à une E.D.S au sens de Stratonovich. Si dans une application (voir 3.1.3) on désire approximer une diffusion donnée sous forme Stratonovich, on peut bien sûr utiliser cette transformation puis appliquer les

méthodes de discrétisation vues jusqu'à maintenant. Mais cette transformation faisant intervenir les dérivées de  $\sigma$ , il peut être préférable d'écrire directement un schéma de discrétisation pour l'E.D.S sous forme Stratonovich.

Citons dans cette direction le schéma F.R.K.S (First order Runge-Kutta method involving the Stratonovich coefficient) de Newton qui est l'analogue du F.R.K.I vu précédemment en définition 10.

**Définition 14** (*Newton[91]*)

Le schéma FRKS est défini par  $\bar{X}_0 = X_0^N$  et :

$$\bar{X}_{t_{k+1}}^N = \bar{X}_{t_k}^N + b(\bar{X}_{t_k}^N)\Delta t + \sigma(\bar{X}_{t_k}^N + \frac{1}{2}\sigma(\bar{X}_{t_k}^N)\Delta W_{k+1})\Delta W_{k+1}. \quad (7.15)$$

Ce schéma a alors les mêmes propriétés que le FRKI vu précédemment. Le schéma de Heun, étudié dans Rumelin[82], peut aussi être utilisé. Il est défini par :

$$\begin{aligned} \bar{X}_{t_{k+1}}^N &= \bar{X}_{t_k}^N + \frac{1}{2}(b(\bar{X}_{t_k}^N) + b(\bar{X}_{t_k}^N + b(\bar{X}_{t_k}^N)\Delta t + \sigma(\bar{X}_{t_k}^N)\Delta W_{k+1}))\Delta t \\ &\quad + \frac{1}{2}(\sigma(\bar{X}_{t_k}^N) + \sigma(\bar{X}_{t_k}^N + b(\bar{X}_{t_k}^N)\Delta t + \sigma(\bar{X}_{t_k}^N)\Delta W_{k+1}))\Delta W_{k+1}. \end{aligned}$$

Ce schéma possède les mêmes ordres de convergence que le schéma de Milshtein.

### 7.3 Approximations en loi.

L'objet d'une telle convergence est en général d'estimer des quantités du type  $\mathbf{E}f(X_T)$  pour une fonctionnelle  $f$ . Quelles en sont les motivations ?

- Le calcul de moments de la diffusion ou autres valeurs moyennes du problème modélisé.
- Faire des estimations en filtrage: Par exemple si

$$Y_t = \int_0^t h(X_s)ds + V_t,$$

avec  $V_t$  un autre brownien indépendant, et si  $Y_t$  est observable, alors estimer  $g(X_T)$  connaissant le passé de  $Y$  jusqu'en  $T$  revient à un calcul du type précédent.

- Résoudre numériquement certaines E.D.P dégénérées: On peut sous certaines hypothèses écrire la solution d'une E.D.P comme l'espérance d'une fonctionnelle d'une diffusion (Karatzas-Shreve[88] par exemple). En grande dimension, ou bien lorsque l'E.D.P est dégénérée, des méthodes probabilistes de type Monte-Carlo peuvent s'avérer pertinentes. Voir 8.4. à ce sujet.

Une des méthodes utilisées pour estimer  $\mathbf{E}f(X_T)$  consiste à approcher cette quantité par  $\mathbf{E}f(\bar{X}_T^N)$  où  $\bar{X}_t^N$  est un schéma de discrétisation de l'E.D.S. On calcule alors en général la deuxième espérance par Monte Carlo, en simulant un grand nombre de trajectoires du schéma discrétisé.

Précisons comment on mesure la vitesse de convergence.

**Définition 15** On dira que le schéma de discrétisation  $\bar{X}_t^N$  converge en loi pour la classe de fonction  $C$  à l'ordre  $\alpha$  si: Pour  $g \in C$ , il existe une constante  $K_T(g)$  telle que:

$$|\mathbf{E}g(X_T) - \mathbf{E}g(\bar{X}_T^N)| \leq K_T(g) \frac{1}{N^\alpha}. \quad (7.16)$$

Bien sûr, si la fonction  $g$  est lipschitzienne, on peut déduire des résultats des parties précédentes la convergence à l'ordre  $\frac{1}{2}$  pour le schéma d'Euler et 1 pour le schéma de Milshtein. Mais en fait la situation est différente: Ces deux schémas convergent à l'ordre 1 et ceci sans condition de commutativité sur  $\sigma$ , comme le montre le résultat suivant.

Notons  $C_{pol}^n$  la classe des fonctions de classe  $C^n$  qui sont, ainsi que leurs dérivées, à croissance polynomiale.

**Proposition 34** (Talay [82][86], Milshtein[78])

Les schémas d'Euler et de Milshtein sont d'ordre 1. Plus précisément, supposons que  $X_0$  ait des moments de tous ordres, alors: Si  $b$  et  $\sigma$  sont de classe  $C_b^4$  alors le schéma d'Euler et de Milshtein sont d'ordre 1 pour la classe des fonctions  $C_{pol}^4$ .

Nous avons vu la preuve de cette propriété pour le schéma d'Euler en 4.5. Précisons que les ordres de convergence peuvent être inférieurs si l'on veut élargir la classe de fonctions  $b$  et  $\sigma$  (voir Mikulevicius-Platen[86]).

Une direction de recherche a été de trouver des schémas ayant un ordre de convergence supérieur à 1. Il en existe plusieurs, nous donnons ici le schéma MCRK de Talay en dimension 1, le cas de la dimension  $d > 1$  est similaire mais beaucoup plus lourd à écrire.

On introduit pour cela les v.a  $\{U_1, V_1, U_2, V_2 \dots\}$  iid telles que

$$\begin{aligned} \mathbf{E}(U_1^p) &= 0 \text{ pour } p = 1, 3, 5, \\ \mathbf{E}(U_1^2) &= 1, \mathbf{E}(U_1^4) = 3, \mathbf{E}(U_1^6) < +\infty. \end{aligned}$$

On peut prendre par exemple

$$\begin{cases} \mathbf{P}(U_1 = \sqrt{3}) = \mathbf{P}(U_1 = -\sqrt{3}) = 1/6 \\ \mathbf{P}(U_1 = 0) = 2/3. \end{cases}$$

**Définition 16** (Talay[83a])

Le schéma MCRK, noté  $(\bar{X}_{t_k})_k$  est défini par  $\bar{X}_0 = X_0$  et

$$\begin{aligned} \bar{X}_{t_{k+1}} = \bar{X}_{t_k} &+ \Delta t(b - \frac{1}{2}\sigma\sigma')(\bar{Y}_{t_k}) + \sqrt{\frac{\Delta t}{2}} \sigma(\bar{X}_{t_k})(U_{k+1} - V_{k+1}) + \sqrt{2\Delta t} \sigma(\bar{Y}_{t_k})V_{k+1} \\ &+ \frac{\Delta t}{2}(\sigma\sigma'(\bar{Y}_{t_k}) - \sigma\sigma'(\bar{X}_{t_k}))V_{k+1}^2 - \frac{\Delta t}{2}\sigma\sigma'(\bar{X}_{t_k})U_{k+1}V_{k+1} \end{aligned} \quad (7.17)$$

avec

$$\bar{Y}_{t_k} = \bar{X}_{t_k} + \frac{\Delta t}{2}(b - \frac{1}{2}\sigma\sigma')(\bar{X}_{t_k}) + \sqrt{\frac{\Delta t}{2}}\sigma(\bar{X}_{t_k})U_{k+1} + \frac{\Delta t}{4}\sigma\sigma'(\bar{X}_{t_k})U_{k+1}^2.$$

Ce schéma converge à l'ordre 2 comme le prouve la

**Proposition 35** (Talay[83a])

Sous les hypothèses  $\mathbf{E}X_0^6 < +\infty$ ,  $b$  et  $\sigma$  de classe  $C_b^6$  et  $f$  de classe  $C_6$  à croissance polynomiale, le schéma MCRK est d'ordre 2.

Remarquons que ce schéma ne fait pas intervenir les dérivés secondes de  $b$  et  $\sigma$  et utilise des v.a qui ne sont pas forcément des gaussiennes et ne converge à priori pas trajectoriellement vers la solution.

Mais est-il toujours pertinent d'utiliser un schéma à l'ordre 2 comme le précédent ? En effet la relative complexité du schéma peut parfois compenser le gain en rapidité de convergence. De plus interviennent dans la définition de ces schémas des dérivées qui peuvent être gênantes dans la pratique.

Il existe à l'inverse un résultat prometteur de Talay et Tubaro qui permet d'avoir des résultats à l'ordre 2 en utilisant simplement le schéma d'Euler. L'idée est de construire des interpolations de type Romberg pour faire un développement limité de l'erreur. On a la

**Proposition 36** (Talay-Tubaro[90])

Notons  $\xi_T = \mathbf{E}f(X_T) - \mathbf{E}f(\bar{X}_T^N)$  l'erreur d'approximation du schéma d'Euler. Si les coefficients  $b$ ,  $\sigma$  et  $f$  sont  $C_b^\infty$ , on peut écrire:

$$\forall k, \quad \xi_T = \frac{C_1}{N} + \dots + \frac{C_k}{N^k} + o\left(\frac{1}{N^{k+1}}\right). \quad (7.18)$$

De plus si

$$Z_T^N = 2\mathbf{E}f(\bar{X}_T^{2N}) - \mathbf{E}f(\bar{X}_T^N),$$

alors  $Z_T^N$  est une approximation à l'ordre 2:

$$|\mathbf{E}f(X_T) - Z_T^N| = O\left(\frac{1}{N^2}\right).$$

On peut aussi, par une seconde interpolation de type Romberg, construire une approximation à un ordre encore supérieur. Des simulations numériques (Talay-Tubaro[90]) semblent montrer le très bon comportement d'une telle méthode par rapport aux méthodes d'ordre 2, du point de vue du temps de calcul, de la précision et même de la stabilité.

Il reste en pratique à calculer  $\mathbf{E}f(\bar{X}_T^N)$ . Cela est fait en général par Monte-Carlo: On simule un grand nombre de réalisations indépendantes du schéma de discrétisation, éventuellement en parallèle, puis on en fait la moyenne. L'inconvénient est que la convergence Monte Carlo est lente, c'est celle du théorème central limite. Des travaux sont en cours pour accélérer cette convergence. Voir 8.3 à ce sujet.

Enfin il y a d'autres travaux dont nous n'avons pas parlé et qui semblent pour le moment moins pertinents pour les applications (ou bien dont nous ne connaissons pas de validation numérique):

- Ecrire des schémas, de type développement de Taylor, ayant un ordre arbitraire de convergence: On peut citer Kloeden-Platen[89][91] et Platen[86]. Il ne semble pas très pertinent en pratique d'avoir des schémas d'ordre supérieur à 2 car ils sont en général très lourds d'utilisation.
- Traiter le cas où les coefficients  $b$  et  $\sigma$  sont seulement holderiens: On peut citer les travaux de Yamada[77], Mikulevicius-Platen[86] pour le schéma d'Euler. Les ordres de convergence sont alors plus faibles.
- Inclure le cas de sauts poissonniens, voir Kloeden-Platen[89].
- Utiliser des estimateurs non biaisés ou des méthodes de réductions de variance, voir Wagner[89], Newton[90c], Chang[87]. C'est une direction de recherche très intéressante mais nous ne connaissons pas pour le moment de résultats numériques confirmants la pertinence de ces méthodes. En particulier Wagner[89] utilise une méthode ingénieuse mais qui nécessite l'évaluation de la densité de transition et semble difficilement utilisable en pratique.

Enfin une autre direction de recherche pourrait être de trouver des résultats de convergence quand la fonction  $f$  est irrégulière. Voir la partie 8.4 à ce sujet.

**Conclusion.** Le schéma d'Euler converge en loi à l'ordre 1 mais permet par interpolation de type Romberg, comme nous l'avons vu, d'obtenir un schéma à l'ordre 2 qui devrait se révéler aussi performant en pratique que les autres schémas d'ordre 2 comme le MCRK de Talay.

Il apparait, pour la convergence en loi, que c'est plutôt la convergence Monte-Carlo du schéma discrétisé qu'il s'agirait d'accélérer pour la pratique, plutôt que de trouver de nouveaux schémas de discrétisation. Voir 8.4 pour quelques compléments.

## 7.4 Autres types de convergence.

Un certain nombres de travaux ne rentrent pas directement dans les deux catégories précédentes. Citons:

- L'approximation d'exposants de Lyapunov.
- L'approximation de la loi invariante d'un processus ergodique.

Décrivons brièvement les résultats connus sur ces deux problèmes.

### 7.4.1 Calcul approché d'exposants de Lyapunov.

La solution d'une E.D.S peut, sous certaines hypothèses de stabilité, converger exponentiellement vers zéro. Nous renvoyons à Arnold[87][74] pour un exposé complet.

On mesure alors cette vitesse de convergence par l'exposant de Lyapunov  $\lambda$  donné par

$$\lambda = \lim_{t \rightarrow \infty} \frac{\log |X_t|}{t} \quad p.s.$$



Pour approximer cet exposant, en suivant Talay[89], l'idée est d'écrire un schéma de discrétisation  $\bar{X}_{t_k}^\delta$  de l'E.D.S avec un pas de discrétisation  $\delta$ , tel que l'exposant de Lyapunov  $\bar{\lambda}^\delta$  du schéma discrétisé vérifie:

$$\exists \alpha > 0, \quad |\bar{\lambda}^\delta - \lambda| = O(\delta^\alpha). \quad (7.19)$$

Puis de calculer alors une approximation  $\bar{\lambda}$  de  $\lambda$  par

$$\bar{\lambda}^\delta \sim \frac{\log |\bar{X}_T^\delta|}{T} \quad (7.20)$$

pour  $T$  assez grand. En fait on utilise une formule un peu différente, en renormalisant les données, pour éviter des phénomènes d'instabilité numérique.

On sait montrer que le schéma d'Euler conduit à une approximation d'ordre  $\alpha = 1$ , on sait aussi construire des schémas d'ordre  $\alpha = 2$  ou plus (voir Talay[89][90a] où un schéma d'ordre 4 est utilisé pour des raisons d'instabilité numérique).

Mais le problème *très délicat* est de savoir comment choisir en pratique le pas de discrétisation  $\delta$  et le temps terminal  $T$  dans (7.17). On ne connaît pas pour le moment de résultats théoriques permettant de faire de tels choix, et des recherches restent à faire dans cette direction.

Nous renvoyons pour plus de renseignements à Talay[90a][89] et aussi Grorud-Talay[90a][90b] pour le cas d'une diffusion définie sur une variété compacte ou sur  $\mathbf{R}^n$ .

Des travaux de Ben Alaya sont en cours sur le sujet.

#### 7.4.2 Approximation de la loi invariante d'un processus ergodique.

Sous certaines hypothèses il existe une unique mesure invariante  $\mu$  et on a pour une fonction  $f$ , par exemple mesurable bornée:

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(X_s) ds = \int f(s) d\mu_s.$$

Une façon d'approximer la loi limite est de remplacer la première intégrale précédente par

$$\frac{1}{N} \sum_{k=1}^N f(\bar{X}_{t_k}^\delta),$$

pour  $N$  assez grand.

Le choix de  $N$  et de  $\delta$  est, là aussi, délicat, On ne connaît pas de résultat général permettant de faire un bon choix pour une précision donnée. Nous renvoyons à Talay[90a][90b] pour un exposé détaillé.

Des travaux de Ben Alaya sont aussi en cours sur le sujet.

## Chapitre 8

### Quelques problèmes ouverts.

L'objet de ce chapitre est de présenter quelques problèmes dont la résolution serait importante pour les applications mais qui n'ont pour le moment que des réponses partielles.

Nous abordons en 8.1 la recherche de schémas de discrétisation, *effectivement simulables* et plus performants que le schéma d'Euler, permettant l'approximation trajectorielle quand la condition de commutativité n'est pas vérifiée. Nous abordons ensuite en 8.2 l'étude de la stabilité des schémas de discrétisation. En 8.3 nous discutons l'utilisation de pas de discrétisation variables pour l'approximation trajectorielle. Enfin nous abordons quelques questions concernant la convergence en loi en 8.4.

## 8.1 Approximation trajectorielle, cas non commutatif.

Considérons une diffusion quelconque dans  $\mathbf{R}^n$  dont nous voulons simuler une approximation trajectorielle. En général cette diffusion ne vérifie pas la condition de commutativité sur  $\sigma$ , rappelée en 7.2.1, et elle ne peut pas à priori se représenter comme une fonctionnelle continue du brownien.

Nous avons vu néanmoins qu'une approximation trajectorielle à des temps discrétisés est possible en utilisant par exemple le schéma d'Euler (Chapitre 4). Nous avons aussi vu que nous ne connaissons pas de schéma facilement simulable (i.e  $P_{\sigma_N}$ -simulable au sens de la définition 9 de 7.2.1) qui permette une convergence plus rapide que le  $1/\sqrt{N}$  du schéma d'Euler. Cependant, en modifiant légèrement la définition classique du schéma de Milshtein, nous avons vu en 5.3 que l'on obtient une approximation en  $1/N$ , que la condition de commutativité soit vérifiée ou non, la contrepartie étant que des intégrales stochastiques multiples apparaissent alors dans la définition du schéma que l'on ne sait pas "bien" simuler. Ces résultats étant par ailleurs expliqués par les travaux de Platen et Wagner rappelés en 7.2.1.

Ainsi nous pourrions améliorer systématiquement la vitesse de convergence du schéma d'Euler si nous savions répondre à la question suivante:

**Problème 8.1:** Est-il possible de simuler conjointement les variables aléatoires:

$$\left( \int_0^\delta W_s^{(i)} dW_s^{(j)} , W_\delta^{(k)} , i, j, k \in [1..d] \right),$$

où  $W_t^{(1)}, \dots, W_t^{(d)}$  sont  $d$  browniens indépendants.

Donnons quelques éléments de réponse. Une première méthode peut être d'approcher les intégrales stochastiques précédentes par une simulation très fine du brownien, ou bien par un développement en série tronqué (voir Kloeden-Platen[89]). Mais, dans les deux cas, l'erreur d'approximation commise est trop élevée. Il y a en fait une difficulté intrinsèque à simuler ces intégrales; elles sont en général des fonctions discontinues de la trajectoire brownienne, i.e non Riemann intégrables. Voir Bouleau[91] à ce sujet.

La bonne solution est de simuler *exactement* ces variables aléatoires. Donnons quelques éléments permettant de penser qu'une telle simulation est possible. Commençons par le cas  $d = 2$  et introduisons l'aire de Lévy

$$S_\delta = \int_0^\delta W_s^{(1)} dW_s^{(2)} - \int_0^\delta W_s^{(2)} dW_s^{(1)}.$$

Simuler le sextuplet du problème 8.1 revient à simuler, d'après la formule d'intégration par partie, le triplet  $(S_\delta, W_\delta^{(1)}, W_\delta^{(2)})$ . Par la formule de Lévy, on connaît la valeur de  $\mathbf{E}(\exp(i\alpha S_\delta) | W_\delta^{(1)} = x, W_\delta^{(2)} = y)$ , il suffit donc de trouver la densité correspondante puis résoudre le problème de simulation.

En suivant Ikeda-Watanabe[81] il est par ailleurs possible de simuler  $S_\delta$  sachant certaines fonctionnelles du brownien.

Pour la dimension quelconque on peut penser utiliser une extension de la formule de Lévy qui apparaît dans Helmes-Schwane[83]. Signalons aussi, pour la dimension 3, qu'une expression de la densité du vecteur constitué par les aires de Lévy deux à deux apparaît dans Helmes[86].

Des travaux menés par T.J. Lyons et J.Gaines sont en cours sur le sujet devant déboucher sur une simulation performante.

## 8.2 Stabilité des schémas de discrétisation.

On peut être surpris par l'absence de résultats sur la stabilité des schémas de discrétisation en comparaison de ce qui existe sur les équations différentielles ordinaires. Une raison en est bien sûr que le cas stochastique est plus difficile. Notons que l'on peut distinguer différents types de stabilité:

- La stabilité *au sens robustesse* concerne la sensibilité du schéma à des imperfections dans sa simulation; imperfections dues à la simulation imparfaite du brownien directeur ou bien encore dues à l'instabilité numériques des calculs.
- La stabilité *au sens de Lyapunov* concerne le comportement de l'erreur d'approximation quand le temps final  $T$  tend vers l'infini.

### Robustesse des schémas d'approximation:

En pratique on commet des erreurs de calcul propres à la machine utilisée, mais aussi des erreurs dans la simulation même des variables aléatoires par l'utilisation de nombres pseudo-aléatoires. Il est donc naturel d'étudier les conséquences de telles erreurs sur les schémas d'approximation.

Une façon de formaliser le problème est d'introduire  $\tilde{X}_t^{u, \Delta t}$  le schéma obtenu en remplaçant, dans la définition du schéma, le brownien par une fonction  $u$ . On étudie alors la convergence de  $\tilde{X}_t^{u, \Delta t}$  vers la solution exacte, quand le pas  $\Delta t$  tend vers zéro et quand la fonction  $u$  tend vers le brownien en un sens à préciser.

Notons qu'une telle étude n'a de sens que si la diffusion peut se représenter comme une fonctionnelle continue du brownien, ce qui est le cas si la condition de commutativité est vérifiée.

Introduisons les notations suivantes:

Soit  $C$  un sous-ensemble de  $C_0([0, T], \mathbf{R}^d)$  contenant les trajectoires browniennes régulières.

$\mathcal{N}_1$  norme sur  $C_0([0, T], \mathbf{R}^d)$  et  $\mathcal{N}_2$  norme sur  $C$ .

$C_\epsilon = \{u \in C, \mathcal{N}_2(u - W) \leq \epsilon\}$ .

Etudier la stabilité ou la robustesse d'un schéma revient à résoudre le problème suivant:

**Problème 8.2:** Pour un schéma d'approximation  $\bar{X}_t^{\Delta t}$ , trouver un quadruplet "optimal"  $(\mathcal{C}, \mathcal{N}_1, \mathcal{N}_2, h)$  où  $h$  est une fonction positive, nulle et continue en 0, tel que

$$\forall \epsilon, \epsilon' > 0, \quad \sup_{u \in \mathcal{C}_\epsilon, \Delta t < \epsilon'} \mathcal{N}_1(X - \bar{X}^{u, \Delta t}) \leq h(\epsilon, \Delta t).$$

On connaît dans ce sens un résultat de stabilité pour le schéma de Milshtein, souvent appelé propriété de robustesse. D'après Talay[83b], on peut prendre pour  $\mathcal{C}$  la classe des fonctions de variation cubique nulle et pour  $\mathcal{N}_1$  et  $\mathcal{N}_2$  la norme de la convergence uniforme. Si on restreint  $\mathcal{C}$  aux fonctions localement holdériennes, alors on peut déterminer explicitement une fonction  $h(x, y)$ .

On trouve parfois dans la littérature l'assertion selon laquelle le schéma d'Euler n'est pas robuste. Il faut tempérer cette assertion. Bien sûr si l'espace  $\mathcal{C}$  est trop large, par le résultat de Wong et Zakai rappelé en 3.1.3, la solution limite est la solution stratonovich, et on ne peut donc pas répondre à la question 8.2.

Comme remarqué par Denis Talay, la classe  $\mathcal{C}$  ne peut être aussi large que celle du schéma de Milshtein. En effet, si l'on reprend la démonstration de la proposition 15 en 4.2, pour que le schéma converge vers la bonne solution, il est nécessaire que le terme  $T_1 = (1 - \sum_i (\Delta u_i)^2)$  converge vers zéro. La classe  $\mathcal{C}$  ne doit donc pas être plus grande que la classe des fonctions de variation quadratique "proche" de celle du brownien. On peut cependant au moins inclure les fonctions de la forme  $u(t) = W_t + \epsilon B_t$  où  $B_t$  est un autre brownien indépendant.

Notons qu'un résultat numérique dans Talay[90a] montre que pour  $\epsilon$  petit le schéma d'Euler converge plus mal vers la solution que le schéma de Milshtien. Il est probable que la dépendance en  $\epsilon$  de la fonction  $h$  soit plus forte pour le schéma d'Euler que pour le schéma de Milshtein.

Nous ne connaissons pas de réponse au problème 8.2 pour le schéma d'Euler. On peut sans doute prendre pour  $\mathcal{C}$  la classe des semi-martingales, muni d'une norme telle que  $\mathcal{C}_\epsilon$  contiennent les semi-martingales de crochet "proche" de celui du brownien.

**Stabilité au sens Lyapunov:** Il existe assez peu de résultats en ce sens, on renvoie à Kloeden-Platen[91]. Les résultats connus sont pour le moment limités au cas linéaire.

Des questions comme l'étude systématique de schémas implicites ou semi-implicites reste à faire, voir Kloeden-Platen[91] pour quelques résultats.

### 8.3 Utilisation de pas variables pour l'approximation trajectorielle.

L'utilisation de méthodes à pas variables peut permettre une simulation plus efficace; du point de vue du temps de calcul, de la place mémoire ou encore de la stabilité numérique.

La méthode présentée dans Gaines-Lyons[91] permet de contrôler l'erreur finale au temps  $T$  par un choix judicieux des pas de discrétisation. Essayons d'en rendre compte en quelques mots. On peut aussi consulter Faure-Gaines[91].

Il s'agit de construire une suite de temps  $t_0 = 0 < t_1 < \dots < t_k < \dots$  telle que le schéma de discrétisation  $\{\bar{X}_{t_k}\}_k$  défini en ces temps possède une erreur finale  $|\bar{X}_T - X_T|$  plus petite qu'un réel  $\epsilon$  fixé par avance.

Définissons pour cela l'erreur locale au  $k^{\text{ième}}$  pas, à un temps  $t \geq t_k$ , par  $el_k(t) = |\bar{X}_t - X_t|$  conditionnellement à  $\bar{X}_{t_k} = X_{t_{k+1}}$ . L'idée est qu'il suffit d'imposer, pour contrôler l'erreur finale, que chaque erreur locale  $el_k(t)$  soit telle que  $el_k(T) < \epsilon$ . Or, en supposant que la propagation d'erreur très petites est linéaire,  $el_k(t)$  vérifie une E.D.S bilinéaire qui, résolue de façon rétrograde, permet de fixer un seuil maximal pour  $el_k(t_{k+1})$  assurant  $el_k(T) < \epsilon$ .

En pratique, on utilise une représentation du brownien sous forme d'un arbre diadique binaire. A l'étape  $k$ , par un développement de Taylor stochastique, il est facile d'approcher  $el_k(t_{k+1})$  par un polynôme simple en  $W_{t_{k+1}} - W_{t_k}$ , et donc de choisir dans l'arbre binaire la profondeur juste suffisante pour que le seuil maximal local de l'étape  $k + 1$  ne soit pas atteint.

Cette méthode semble bien fonctionner sur les exemples que nous avons pu voir, problèmes d'estimations statistiques ou d'étude de flots, et semble être très performante en temps cpu par l'utilisation efficace et "optimale" de l'arbre brownien. Nous savons que les auteurs ont une preuve de la convergence de cet algorithme pour les schémas classiques, sauf pour le schéma d'Euler. Il est certain qu'une telle méthode devrait être amenée à se développer.

Une autre direction pourrait être d'utiliser une *interpolation de type Romberg* pour l'approximation trajectorielle, dans le même esprit que les travaux de Talay-Tubaro rappelés en 7.3, mais trajectoire par trajectoire. Pour le schéma d'Euler classique il conviendrait tout d'abord de connaître la vitesse exacte de convergence par un théorème de type logarithme itéré. En première approximation, si le schéma est d'ordre  $\alpha$  au sens de (7.13), alors une simulation du schéma avec un pas de  $\Delta t$  puis avec un pas de  $2\Delta t$  selon la même trajectoire brownienne, devrait donner un schéma  $Z_t^{\Delta t} = 2^\alpha \bar{X}_t^{\Delta t} - \bar{X}_t^{2\Delta t}$  d'ordre de convergence supérieur. Sur les quelques simulations que nous avons effectué, cette méthode semble donner une approximation plus stable quand le temps croît.

On peut aussi penser faire une interpolation similaire sur le schéma E.C.B du chapitre 6 qui possède un ordre de convergence un peu meilleur (Proposition 16 et 28).

Un autre problème concerne la *simulation très fine* des diffusions permettant l'étude de propriétés locales comme les points de dégénérescence. Pour cela il doit être possible d'étendre les résultats du chapitre 6, par exemple pour simuler un schéma de type E.C.D en dimension strictement plus grande que 1. Un tel schéma semble naturellement "proche" de la diffusion correspondante en ce sens qu'il doit posséder beaucoup plus de propriétés communes avec la diffusion exacte, que les schémas classiques à pas constants.

## 8.4 Sur l'approximation en loi

Il est souvent avancé, pour motiver l'étude de la convergence Monte Carlo des schémas de discrétisation, l'application possible à la résolution de certaines E.D.P. La pertinence d'une telle application n'est cependant pas clair.. Commençons par discuter deux exemples simples.

Considérons l'opérateur

$$\mathcal{L} = \sum_i b_i \partial_{x_i} + \frac{1}{2} \sum_{i,j} a_{i,j} \partial_{x_i x_j}^2,$$

où  $a(x)$  est définie positive et s'écrit  $a(x) = \sigma(x)\sigma^*(x)$ . La solution du système

$$\begin{cases} \frac{\partial v}{\partial s} + \mathcal{L}v + \alpha v = 0 & \text{sur } [0, T] \\ v(T, x) = f(x), \end{cases}$$

peut se représenter par

$$v(s, x) = \mathbf{E}(f(X_T^{s,x}) \exp(\int_s^T \alpha(u, X_u^{s,x}) du))$$

où  $X_T^{s,x}$  est la solution de l'E.D.S associée à l'opérateur  $\mathcal{L}$  partant de  $x$  au temps  $s$ . On peut donc, par une approximation de type Monte Carlo, connaître la solution  $v(s, x)$  par simulation numérique. Cependant un très grand nombre de simulations indépendantes sont nécessaires pour connaître la solution au seul point  $x$ . Une telle méthode risque donc de n'être pertinente que par l'utilisation d'une machine fortement parallèle, et dans des problèmes particuliers comme dans les très grandes dimensions où les méthodes traditionnelles d'analyse numérique marchent mal.

La solution est plus complexe dans le cas suivant. Considérons le problème de Poisson:

$$\begin{cases} \frac{1}{2} \Delta u = -g & \forall x \in D \\ u = f & \forall x \in \partial D \end{cases}$$

Si l'ouvert  $U$  est suffisamment régulier et si les fonctions  $f$  et  $g$  sont suffisamment régulières, alors la solution se représente sous la forme

$$u(x) = \mathbf{E}^x(f(W_{\tau_D}) + \int_0^{\tau_D} g(W_t) dt),$$

où  $\tau_D = \inf\{t > 0, W_t \notin D\}$ .

La situation est ici plus complexe car il faut simuler  $(W_{\tau_D}, \tau_D)$ . Pour un tel exemple on peut penser faire une simulation spatiale du brownien, dans l'esprit du chapitre 2, pour obtenir une "bonne" approximation de la solution exacte par Monte Carlo.

Dans le cas plus général où l'opérateur n'est pas le Laplacien vient s'ajouter deux nouveaux problèmes:

- Le premier est de savoir simuler la diffusion jusqu'à sa sortie de l'ouvert. En effet, la convergence de type Monte Carlo des schémas de discrétisation est toujours faite sur un intervalle de temps borné, ce qui n'est plus ici le cas.
- Le deuxième concerne la régularité de la fonction  $f$ . En effet le cas le plus intéressant, i.e le cas où les méthodes traditionnelles marchent mal, est celui où la fonction  $f$  est peu régulière. Or justement un tel cas n'est pour le moment pas couvert par les résultats sur la convergence Monte Carlo (voir 7.3).

L'utilisation de schémas de discrétisation spatiale comme au chapitre 6, ou encore l'utilisation de méthodes mixtes comme utilisé en 6.4, pourraient peut-être permettre de résoudre ce genre de problèmes. Des travaux que nous menons avec Martin Clark sont en cours sur le sujet.

Signalons pour mémoire qu'il existe d'autres méthodes probabilistes pour résoudre une classe plus large d'E.P.D, comme les équations de réaction-diffusion, et connues sous le nom de méthodes particulières. Voir par exemple Bernard[91] pour un panorama de ces méthodes.

Un autre problème important pour les applications consiste à accélérer la convergence Monte Carlo, par des méthodes de réductions de variance par exemple. Des travaux prometteurs, menés par Nigel Newton, sont en cours sur le sujet. Une utilisation astucieuse de la méthode du shift de Bernoulli rappelée en 2.5.3, pourrait aussi permettre accélérer la convergence Monte Carlo.

Enfin signalons un travail prometteur de Ogawa[91] sur l'approximation des diffusions non-linéaires au sens de McKean. Pour simuler une approximation trajectorielle d'une telle diffusion, une estimation Monte Carlo du schéma est nécessaire à chaque étape. Nous ne connaissons pas encore de résultats numériques confirmant la pertinence d'une telle méthode.







# Bibliographie



Arnold L.[87], Lyapunov Exponents of Non linear Stochastic System, *Nonlinear Stochastic Dynamic Engineering Systems*, G.I. Schuelle & F. Ziegler (Eds.), Proceedings of the IUTAM Symposium, Innsbruck, Springer-Verlag, 1987.

Arnold L.[74], *Stochastic Differential Equations*, Wiley, New-York, 1974.

Bauer H.[66], Harmonische Räume und ihre Potential theorie, *Lect. Notes in Maths.* 22, Springer-Verlag, 1966.

Bichteler K.[81], Stochastic Integration and  $L^p$  Theory of Semimartingales, *The Annals of Probability*, Vol 9, No 1, 49-89, 1981.

Ben Alaya M.[91], Sur le calcul approché d'espérances de variables aléatoires dépendant d'un temps d'arrêt, Preprint LAMM, soumis pour publication, 1991. Voir aussi l'article dans le recueil sur les probabilités numériques dans la collection INRIA, 1991.

Bernard P.[91], Les méthodes particulières stochastiques: Une brève présentation, à paraître dans un recueil sur les probabilités numériques dans la collection INRIA, 1991.

Bouleau N.[91], Irregular and Simulatable Functionals on Wiener Space, à paraître dans un recueil sur les probabilités numériques dans la collection INRIA, 1991.

Bouleau N.[90a], *Calculs numériques en Probabilité*, Cours de D.E.A 1991 de Probabilité de Paris VI.

Bouleau N.[90b], On effective computation of expectations in large or infinite dimension, *Journal of Computational and Applied Mathematics* 31 , 23-34, 1990.

Bouleau N.[86], *Probabilités de l'ingénieur, variables aléatoires et simulation*, Hermann, Paris, 1986.

Chang, C.C.[87], Numerical solution of Stochastic Differential Equations with constant diffusion coefficients, *Math. Computat.* 49, 523-542, 1987.

Clark J.M.C.[88], The discretization of Stochastic Differential Equations: A primer, in H. Neunzert (ed.) *Road vehicle Systems and Related Mathematics, Proc. Second Workshop Torino 1987*, series Mathematical Methods in technology, Tenbres Reided, 1988.

Clark J.M.C.[84], Asymptotically optimal quadrature formulae for stochastic integrals, *Proceedings of the 23rd Conference on Decision and Control*, Las Vegas, Nevada, USA, IEEE, 1984, 712-715.

Clark J.M.C.[82], An efficient approximation for a class of stochastic differential equations, *Advances in Filtering and Optimal Stochastic Control*, eds. W.H Fleming and L.G Gorostiza. Proceedings of IFIP-WG7/1 Working Conference, Cocoyoc, Mexico, 1982. *Lecture notes in control and information sciences* 42, Springer-Verlag, 1982.

Clark J.M.C.[80], An efficient Approximation Scheme for a class of Stochastic Differential Equation of Nonlinear Filtering, *Communication Systems and Random Process Theory*, J. Skwirzynskii (ed.), Sijthoff & Noordhoff, Alphen aan den Rijn, 1980.

Clark J.M.C. - Cameron, R.J.[80], The Maximum rate of Convergence of Discrete Approximations for Stochastic Differential Equations, *Stochastic Differential Systems, B. Grigelionis (ed.) Lecture Notes in Control and Information Sciences*, 25, Springer-Verlag, Berlin, 1980.

Devroye.L.[86], *Non uniform Random Variate Generation*, Springer-Verlag, 1986.

Doss H.[77], Liens entre Equations Différentielles Stochastiques et Ordinaires, *Ann. Inst. H. Poincaré* 13. 99 , 1977.

Faure O. - Gaine J.[91], Simulation trajectorielle des diffusions, à paraître dans un recueil sur les probabilités numériques dans la collection INRIA, 1991.

Faure O.[90], Numerical Pathwise Approximation of Stochastic Differential Equations, Preprint LAMM 90/7, 1990. A paraître dans *Applied Stochastic Models and Data Analysis*, 1992.

Feller[71], *An introduction to probability theory and its applications*, Vol II, second edition, John Wiley-Sons, 1971.

Friedman, A.[75], *Stochastic Differential Equations and Applications*, Vol. I, Academic Press, New-York, 1975.

Friedman, A.[76], *Stochastic Differential Equations and Applications*, Vol. II, Academic Press, New-York, 1976.

Gaines J. - Lyons T.J.[91], Variable Step Size Control in the Numerical Solution of Stochastic Differential Equations, soumis pour publication, 1991.

Gard T.C.[88], *Introduction to stochastic differential equations*, M Dekker, 1988.

Geraldi-Marchetti-Rosa[84], Simulations of diffusion with boundary conditions, *Systems & Control letters* 4, No 5, 253-261, 1984.

Gikhmann I.I - Skorokhod A.V.[79], *The Theory of Stochastic Processes*, Vol I-III, Springer-Verlag, 1979.

Gorod A., Talay D.[90a], Approximation of Lyapunov Exponents of Stochastic Differential Systems on compact Manifolds, *Proc. 9<sup>th</sup> conf. Analysis & Optimization of Systems*, Springer-Verlag, 1990.

Gorod A., Talay D.[90b], Approximation of Lyapunov Exponents of Stochastic Differential Systems on  $\mathbf{R}^n$ , Preprint, 1990.

Helmes K.[86], On Levy's area processus. *Stochastic Differential Systems, Proceedings of the 3rd Bad Honnel Conference*, 1985, *Lect. Notes in Control and Inf. Sciences*, 78. Springer-Verlag, Heidelberg, 1986.

Helmes K. - Schwane A.[83], Lévy stochastic area formula in higher dimensions, *J. Fct. Analysis* 54, 177-192, 1983.

Ikeda N. - Watanabe S.[81], *Stochastic Differential Equations and Diffusion Processes*, North-

Holland, Amsterdam (Kodansha Ltd., Tokyo), 1981.

Imhof J.P.[84], Density factorization for Brownian Motion, meandre and the three dimensionnel Bessel processes and Applications, *J. Appl. Prob.*, 21, 500-510, 1984.

Itô K. - McKean H.P.[65], *Diffusion processes and their sample paths*, Springer-Verlag, 1965.

Jacod J. - Shiryaev A.N.[87], *Limit Theorems for Stochastic processes*, Springer-Verlag, Berlin, 1987.

Kallianpur B.[80], *Stochastic filtering theory*, Springer, Berlin, 1980.

Kanagawa S.[88], On the rate of convergence for maruyama's approximate solutions of S.D.E, *Yokohama Mathematical Journal*, Vol 36, 1988.

Karatzas I. - Shreve S.E.[88], *Brownian Motion and Stochastic Calculus*, Springer-Verlag, 1988.

Kent[78], Some probabilistic properties of Bessel functions, *The Annals of Prob.*, Vol. 6, No 5, 760-770, 1978.

Kloeden P.E. - Platen E.[91], *The Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, 1991.

Kloeden P.E. - Platen E.[89], A Survey of Numerical Methods for Stochastic Differential Equations, *Journal of Stochastic Hydrology and Hydraulics*, 155-178, 1989.

Knight F.B.[81], *Essential of Brownian Motion and Diffusions*, Math. Surveys, Vol. 18, American Mathematical Society, Providence, R.I., 1981.

Kree-Soize[83], *Mécanique Aléatoire*, Dunod, Paris, 1983.

Krylov N.V.[80], *Controlled Diffusion Processes*, Springer-Verlag, 1980.

Kubilius K.Q. - Platen E.[90], Rate of convergence of the Euler Approximation for Diffusion Processes with jumps, preprint P-Math 28/90, Karl-Weierstress institut fur Mathematik, Berlin, 1990.

Kunita H.[82], Stochastic Differential Equations and Stochastic Flows of Diffeomorphisms, *Cours de l'école d'été de Probabilité de Saint-Flour, 1982, Lect. Notes in Math.*, Vol 1097, Springer-Verlag, 1982.

Knuth D.E.[81], *The Art of Computer Programming*, 2nd ed. Addition-Wesley, Reading, Mass.

Kushner[77], *Probability Methods for Application in Stochastic Control and for Elliptic Equations*, Academic Press, London, 1977.

Kushner[74], On the weak convergence of interpolated Markov chains to a diffusion, *Ann. Prob.*, 40-50, 1974.

Lamberton D. - Lapeyre B.[91], *Cours de Mathématiques financières de l'ENPC*, à paraître chez SIAM.

Langevin P.[1908], Sur la théorie du mouvement brownien, *C.R. Acad. Sci. Paris*, 146, 530-533, 1908.

Leblong J., Talay D.[86], Simulation of Diffusion Process with Presto Building Systems like Presto with Adadgio, *Proc. French-Belgian Stat. Congress*, Nov. 1986, Rouen, Edited by J.P. Raoult.

Le Gall J.F.[90], Introduction au mouvement brownien, exposé introductif au congrès de la SMF, 1990.

Lepingle D. - Ribemont B.[89], Un schéma multipas d'approximation de l'équation de Langevin, Preprint, 1989.

Liske - Platen E.[87], Simulation studies on time discrete diffusion approximations, *Math. Comput. Simulation* 29, No 3-4, 253-260, 1987.

Liske - Platen E. - Wagner W.[82], About mixed multiple Wiener integrals, Preprint P-Maths 23/82 I Math., Adw der DDR, Berlin, 1982.

MacShane E.J.[76], The Choice of a Stochastic Model for a Noise System, *Math. Prog. Stud.* 6, 79-92, 1976.

MacShane E.J.[74], *Stochastic Calculus and Stochastic models*, Academic Press, New-york, 1974.

Marsaglia G. - Zaman A.[91], A new class of random number generators, *The Annals of Applied Probability*, Vol. 1, No. 3, 462-480, 1991.

Maruyama G.[55], Continuous Markov processes and stochastic equations, *Rend. Circ. Mat. Palermo* 4, 48-90, 1955.

Mikulevicius R. - Platen E.[88], Time Discrete Taylor Approximation for Itô Processes with jump componeent, *Math. Nachr.*, 138, 93-104, 1988.

Mikulevicius R. - Platen E.[86], Rate of convergence of the Euler approximation for diffusion processes, Preprint P-MATH-38/86, Karl Weierstrass Institut fur Mathematik der Akademie der Wissenschaften der DDR, Berlin, 1986.

Milshtein G.N.[87], A Theorem of the Order of Convergence of Mean Square Approximation of Solutions of system of Stochastic Differential Equations, *Teor. Veroyatnosti Priuenen* 32, No 4, 809-811, 1987.

Milshtein G.N.[85], Weak Approximation of Solution of system of Stochastic Differential Equations, *Theory Prob. Appl*, Vol XXX, No 4, 750-766, 1985.

Milshtein G.N.[78], A Method of Second Order Accuracy Integration of Stochastic Differential Equations, *Theory Prob. Appl* 23, 396, 1976.

Milshtein G.N.[74], Approximate Integration of Stochastic Differential Equations, *Theory Prob. Appl* 19, 557, 1974.

Newton N.J.[91], Asymptotically Efficient Runge-Kutta Methods for a Class of Itô and Stratanovich Equations, *SIAM J. of Applied Mathematics*, Vol. 51. 2, 542-567, 1991.

Newton N.J.[90a], An Efficient Approximation for Stochastic Differential Equations on the Partition of Symmetrical First Passage Times, *Stochastic and St. Reports* 29, 227-258, 1990.

Newton N.J.[86], An asymptotically Efficient Difference Formula for solving Stochastic Differential Equations, *Stochastics* 19, 1986, 175-206.

Oksendal B.[85], *Stochastic Differential Equations an Introduction with Applications*, Springer-Verlag, Berlin Heidelber, 1985.

Ogawa S.[91], Monte Carlo simulation of nonlinear diffusion processes, soumis pour publication, 1991.

Pages G.[86], Un théorème de convergence fonctionnelle pour les semi martingales, *Séminaire de Probabilité XX, Lect. Notes in Math.*, 1204, 572-611, Springer-Verlag, 1986.

Pages G.[89], Continuité de Skorokhod des quantités liées aux temps d'entrée dans un fermé, Communication personnelle non publiée.

Pardoux E. - Talay D.[85], Discretization and Simulation of Stochastic Differential Equations, *Acta Applicandae Mathematicae* 3, 23-47, 1985.

Pei Hsu[85], Brownian exit distribution for a ball, *Seminar on Stochastic Processes*, Birhkauser, Boston, 1985.

Platen E.[86], Derivative Free Numerical Methods for Stochastic Differential Equations, *Lect. Notes in Control and Information Sciences*, 95, 1986.

Platen E. - Rebolledo[85], Weak Convergence of Semimartingales and DSiscretisation Methods, *Stochastic Processes & Applications* 20, 41-58, 1985.

Platen E.[84], On first exit times of diffusions, *Stoch. Diff. Systems 1984, Marseille Luminy, 192-195, Lect. Notes in Control and Infor. Sci.* 69, Springer-Verlag, 1985.

Platen E.[83], Approximation of first exit times of diffusion and approximate solution of parabolic equation, *Math Nachr.* III, 127-164, 1983.

Platen E.[82], An Appromation Method for a Class of Itô processes with jump component, *Liet. Mat. Rink.* 22, 124-136, 1982.

Platen E. Wagner W.[82], On a Taylor Formula for a class of Ito processes, *Probability Math. Stat.*, 3, 37-51, 1982.

Platen E.[81], An Appromation Method for a Class of Itô processes, *Liet. Mat. Rink.*, Vol. XXI, 121-133, 1981.

Platen E.- Wagner W.[78], Approximation of Itô Integral Equations, *Proc IFIP-WG 7/1 Working Conference*, Vilnius, 1978(B. Grigelionis ed.), Stochastic Differential Systems, Springer-Verlag, Berlin, 172-176, 1980.

Rogers L.C.G -Williams D.[87], *Diffusion, Markov Processes, and Martingales*, John Wiley and Sons, Chichester, 1987.



Rumelin W.[82], Numerical Treatment of Stochastic Differential Equations, *Siam J. Num. Anal.* 19, 604-613, 1982.

Sussmann H.J.[78], On the Gap Between Determinist and Stochastic Ordinary Differential Equations, *Ann. Prob.* 6, 19, 1978.

Talay D.[90a], Simulation and Numerical Analysis of Stochastic Differential Systems: A Review, *Rapport de Recherche INRIA n° 1313*, 1990. Soumis pour publication.

Talay D.[90b], Second Order Discretization Schemes of Stochastic Differential Systems for the computation of the invariant law, *Stochastics and Stochastics Reports*, 29(1), 13-36, 1990.

Talay D. - Tubaro L.[90], Expansion of the global error for numerical schemes solving Stochastic Differential Equations, *Rapport INRIA 1990*.

Talay D.[89], Approximation of upper Lyapunov exponents of bilinear stochastic differential systems, *Rapport INRIA 965*, 1989. A paraître dans SIAM Journal on Numerical Analysis.

Talay D.[86], Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution, *Mathematical Modelling and Numerical Analysis*, 20, 141-179, 1986.

Talay D.[83a], Efficient numerical schemes for the approximation of expectations of functionals of S.D.E., *Filtrating and Control of Random Processes*, H.Korezlioglu & G.Mazzioti & J. Szpirglas (Eds.), *Proceedings of the ENST-CNET Colloquium, Paris, 1983, Lecture Notes in Control and Information Sciences* 61, Springer-Verlag, 1984.

Talay D.[83b], Résolution Trajectorielle et Analyse Numérique des Equations Differentielles Stochastiques, *Stochastics* 9, 275-306, 1983.

Talay D.[82], Analyse Numérique des Equations Différentielles Stochastiques, *Thèse de 3<sup>e</sup> Cycle, Univ. de Provence*, 1982.

Unny T.E.[84], Numerical integration of stochastic differential equations in catchment modeling, *Water Resources Res.* 20, 360-368, 1984.

Van Kampen N.G.[81], *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, 1981.

Wagner W. [89], Unbiased Monte Carlo Estimators for functionals of weak solution of Stochastic Differential Equations, *Stochastics and St. Reports* 28, 1-20, 1989.

Wendel J.G.[80], Hitting Sphere with Brownian Motion, *The Annals of Prob.*, vol. 8, No 1, 164-169, 1980.

Wong E. - Zakai M.[65], On the convergence of ordinary integral to stochastic integrals, *Ann. Math. Statist.*, 36, 1560-1564, 1965.

Wright D.J.[74], The digital simulation of Stochastic Differential Equations, *IEEE Trans. Automat. Control*, 19, 75-76, 1974.

Yamada T.[77], Sur une construction des solutions d'Equations Différentielle Stochastiques dans

le cas non lipschitzien, *Sem Prob XIII*, 114-131, Springer-Verlag, 1977.

Yor M.[80], Loi de l'indice du lacet brownien et distribution de Hartman-Watson, *Zeitschrift für Wahr.*, 53, 71-95, 1980.

“



## **Annexe I: Générateur pseudo-aléatoire et simulation du hasard.**



## Générateurs pseudo-aléatoires:

Simuler une suite de nombres au hasard, c'est à dire simuler les réalisations successives d'une suite de variables aléatoires  $U_1, U_2, \dots, U_n, \dots$  iid uniformes sur  $[0, 1]$  peut apparaitre comme une gageuse. On sait en effet, depuis un résultat assez récent de Martin-Löf, que, en caricaturant, aucun algorithme ne peut générer une suite au hasard !

On sait néanmoins très bien simuler des suites dites *pseudo-aléatoires* qui se comportent "presque" comme de vraies suites aléatoires. Par des considérations théoriques, mais aussi par des tests statistiques expérimentaux, on vérifie que ces suites ont de *bonnes* propriétés d'équirépartition, de non corrélation ... Les générateurs de nombres pseudo-aléatoires les plus répandus sont basés sur des méthodes de congruence. On simule une suite pseudo-aléatoire  $x_1, x_2, \dots, x_n, \dots$  sur  $[0, m]$ , par

$$x_{n+1} = ax_n + b \mod m$$

où les entiers  $a$ ,  $b$  et  $m$  sont judicieusement choisis. On obtient alors une suite pseudo-aléatoire sur  $[0, 1]$  en divisant chaque terme de la suite par  $m$ . Bien sûr le choix de  $a$ ,  $b$  et  $m$  est affaire de métier, et on renvoie à Knuth[81] pour plus de renseignements.

Signalons que d'autres méthodes existent qui permettent d'obtenir des générateurs de période aussi grande que  $2^{1376}$  de façon très efficace. Voir Marsaglia-Zaman[91] à ce sujet qui présente un bref panorama des méthodes connues.

Enfin précisons que, bien que l'on trouve sur tout ordinateur au moins un générateur de nombres pseudo-aléatoires, il est recommandé de les utiliser avec prudence. Par exemple en leur faisant tout d'abord passer quelques tests statistiques, ou encore en cherchant leur période. Nous avons préféré implémenter un générateur par nous même. De nombreux arguments plaident en faveur d'une telle attitude. Il existe dans la littérature des générateurs qui ont fait leurs preuves; ils n'ont pas, au fil des années, révélé de défauts flagrants. De plus les résultats sont reproductibles: Connaissant le générateur utilisé et le point de départ de la suite, on peut reproduire la même suite pseudo-aléatoire et ainsi les mêmes résultats. La période de la suite est en général facile à déterminer. Enfin les programmes sont plus portables. Le générateur utilisé tout au long de cette thèse, qui provient de Knuth[81], est basé sur la congruence suivante:

$$x_{n+1} = 31\,415\,821\,x_n + 1 \mod 100\,000\,000,$$

le point de départ étant 1 234 567. Sa période est de 100 000 000.

Il est implémenté de façon à éviter les erreurs d'arrondis dans les divisions. Voici le source en C:

```

int sed_random(xn)

int xn;
{
/* Calcul de  $xn := a \cdot xn + 1 \bmod m$  */

int p1,p0,q1,q0,mult; /* Pour la fonction Mult */
int m,m1,a;           /* Constantes du generateur */
int toto,tutu;        /* Intermediaires */

m = 1000000000;
m1 = 10000;
a = 31415821;

p1  = xn / m1;          /*p1  = [xn/m1] */
p0  = xn - m1 * p1;     /*p0  = xn mod m1 */
q1  = a / m1;           /*q1  = [a/m1] */
q0  = a - m1 * q1;      /*q0  = a mod m1 */
toto = (p0*q1 + p1*q0) / m1;
tutu = (p0*q1+p1*q0 - toto*m1)*m1+p0*q0;
toto = tutu / m;
mult = tutu - toto * m; /*mult=(p0q1+p1q0 mod m1)*m1 + p0q0*/
tutu = mult + 1;
toto = tutu / m;
xn   = tutu - toto * m; /*xn   = (mult+1) mod m*/
return (xn);
}

```

**Performance du générateur:** Quant nous avons indiqué dans cette thèse un temps *cpu* de simulation, ce temps dépend bien évidemment de la machine utilisée, en général ici un SPARC 2. Nous donnons aussi la correspondance en temps *svu*, dont l'unité est le temps de Simulation d'une Variable Uniforme sur  $[0, 1]$  par notre générateur. La correspondance est ici:  $1 \text{ svu} \simeq 3 \cdot 10^{-5} \text{ cpu}$ . Il est ainsi plus aisé de faire des comparaisons de performances indépendamment des machines utilisées.

### Génération de variables aléatoires:

On admet une fois pour toute que l'on sait simuler une suite iid uniforme sur  $[0, 1]$ , notée  $U_1, U_2, \dots, U_n, \dots$ . Les deux méthodes les plus utilisées pour générer une variable aléatoire sont les méthodes d'inversion de la fonction de répartition et la méthode de rejet. Nous renvoyons à Bouleau[86] pour une description de ces méthodes, ainsi que leur utilisation pour simuler les variables aléatoires courantes. Donnons le cas d'une gaussienne:

Deux réalisations successives  $N_1, N_2$  d'une gaussienne centrée réduite sont simulées par

$$N_1 = (-2\text{Log}U_1)^{1/2} \cos(2\pi U_2)$$

$$N_2 = (-2\text{Log}U_1)^{1/2} \sin(2\pi U_2).$$

Il arrive bien souvent que ces méthodes ne suffisent pas. Au chapitre 2 par exemple nous avons utilisé un algorithme de simulation quand la densité, qui ne possède pas d'expression simple, est donnée par une série. On renvoie à Devroye[86] pour un exposé très complet sur les différentes méthodes de simulation des variables aléatoires.